Statistisk sentralbyrå
Statistics Norway

Can the child penalty be reduced?

Evaluating multiple policy interventions

Martin Eckhoff Andresen og Emily Nix

TALL SOM FORTELLER
DISCUSSION PAPERS
983

Discussion Papers No. 983, July 2022
Statistics Norway, Research Department

Martin Eckhoff Andresen og Emily Nix

Can the child penalty be reduced?
Evaluating multiple policy interventions

Abstract:
Children cause large earnings drops for mothers but not fathers, a stylized fact known as the "child penalty" that explains a substantial portion of remaining gender income gaps. Can policy reduce the child penalty? We first document how changes in the child penalty over a long time horizon in Norway correlate with major family policy reforms. Next, we evaluate two possible interventions: paternity leave and high-quality childcare. We find no impact of paternity leave on child penalties or a measure of father's preferences for childcare. In contrast, a year of publicly provided childcare reduces child penalties by 23% during the years of use. These results suggest governments can act to reduce child penalties, but providing alternatives to the mother's time, such as quality childcare, is more effective than paternity leave.

Keywords: Gender wage gap, child penalty, paternity leave, childcare

JEL classification: J21, J23, J22, J71

Address: Martin Eckhoff Andresen, Statistics Norway, Research Department. E-mail: mrt@ssb.no

Acknowledgements: We thank seminar participants at ASU, Claremont McKenna, CSU Fullerton, Duke, EALE, Erasmus, LSU, Purdue, RAND, SOLE, SOFI Stockholm University, Stanford, Statistics Norway, UC Riverside, University of Oslo, University of Rochester, VATT Helsinki, Aarhus University and Warwick University. We also thank Heather Antecol, Manuel Bagues, Sebastian Calonico, Matias D. Cattaneo, Nina Drange, James Fenske, Yana Gallen, Trude Gunnes, Andrea Ichino, Edwin Leuven, Petra Persson, Adam Sheridan, Thor Olav Thoresen, Kenneth Aarskaug Wiik, Natalia Zinovyeva, and Antonio Dalla Zuanna for helpful comments and suggestions. All errors remain our own. Andresen gratefully acknowledges financial support from the Norwegian Research Council (grant no. 236947). This paper was previously circulated as "What Causes the Child Penalty and How Can it be Reduced? Evidence from Same-Sex Couples and Policy Reforms". The first part of the previously circulated draft is now "What Causes the Child Penalty? Evidence from Same Sex and Adopting couples".

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

At barn henger sammen med store fall i inntekten til mødre, men ikke fedre, er et tilbakevendende mønster i en rekke land som på engelsk betegnes som «child penalty». Over tid har andre forklaringer på forskjellen mellom menns og kvinners inntekter, slik som utdanning, blitt mindre, slik at ankomsten av barn i dag kan forklare store deler av den gjenværende kjønnsforskjellen i inntekt (Kleven mfl. 2019).

I denne artikkelen spør vi forvidt familiepolitikk kan bidra til å redusere slike «child penalties». Vi dokumenterer først hvordan ankomsten av barn påvirker menn og kvinners inntekt forskjellig over en tidshorisont fra 1970 til i dag. Videre undersøker vi hvordan endringer i dette målet korrelerer med endringer i familiepolitikk, slik som innføring og forlengelse av foreldrepermisjon og utvidelse av barnehagetilbudet.

Dette motiverer undersøkelsen av hvordan to nyere politikktiltak bidrar til å redusere «child penalties», nemlig (utvidelsen av) fedrekvoten i foreldrepermisjonen og utvidelsen av tilgangen på subsidiert barnehage for de minste barna tidlig på 2000-tallet. For å evaluere effekten av fedrekvoten bruker vi utvidelsene av denne som ble innført 1. juli 2005, 2006, 2009, 2011, 2013 og 2014, og utnytter at par som fikk barn like før 1. juli var omfattet av den gamle fedrekvoten mens par som fikk barn like etter var omfattet av den nye. Vi kombinerer alle reformene for å maksimere statistisk kraft til å avdekke effekter på begge foreldres inntekt, men finner små og insignifikante effekter. Dette innebærer at fedrekvoten har lite potensiale for å gjøre noe med den ulike fordelingen av kostnadene ved å ha barn som målt ved redusert inntekt. Vi undersøker også et mål på normer, nemlig i hvilken grad fedre velger å ta ut mer pappaperm ved neste fødsel dersom reformene utsatte dem for mer pappaperm, og finner lite tegn på at fedres preferanser til å ta del i permisjonen blir påvirket av pappakvotene.

For å evaluere effekten av subsidiert barnehage for de minste barna bruker vi den store utrullingen av barnehager for 1- og 2- åringer som fulgte barnehageforliket i 2002. Som tidligere dokumentert av Andresen og Havnes (2019) ble nye barnehager rullet ut over tid og mellom kommuner på en måte som ser ut til å være så godt som uavhengig av trender i arbeidstilbud. Resultatene tyder på at barnehagebruk reduserer «child penalties» med omtrent 23% per år i de årene tidlig barnehage benyttes, men at det ikke er tegn på langsiktige virkninger utover dette. Oppsummert kan altså den kjønnede fordelingen av kostnader ved barn påvirkes av politikk, men subsidiert barnehage later til å være et mer effektivt virkemiddel enn fedrekvoter i foreldrepermisjon.

1 Introduction

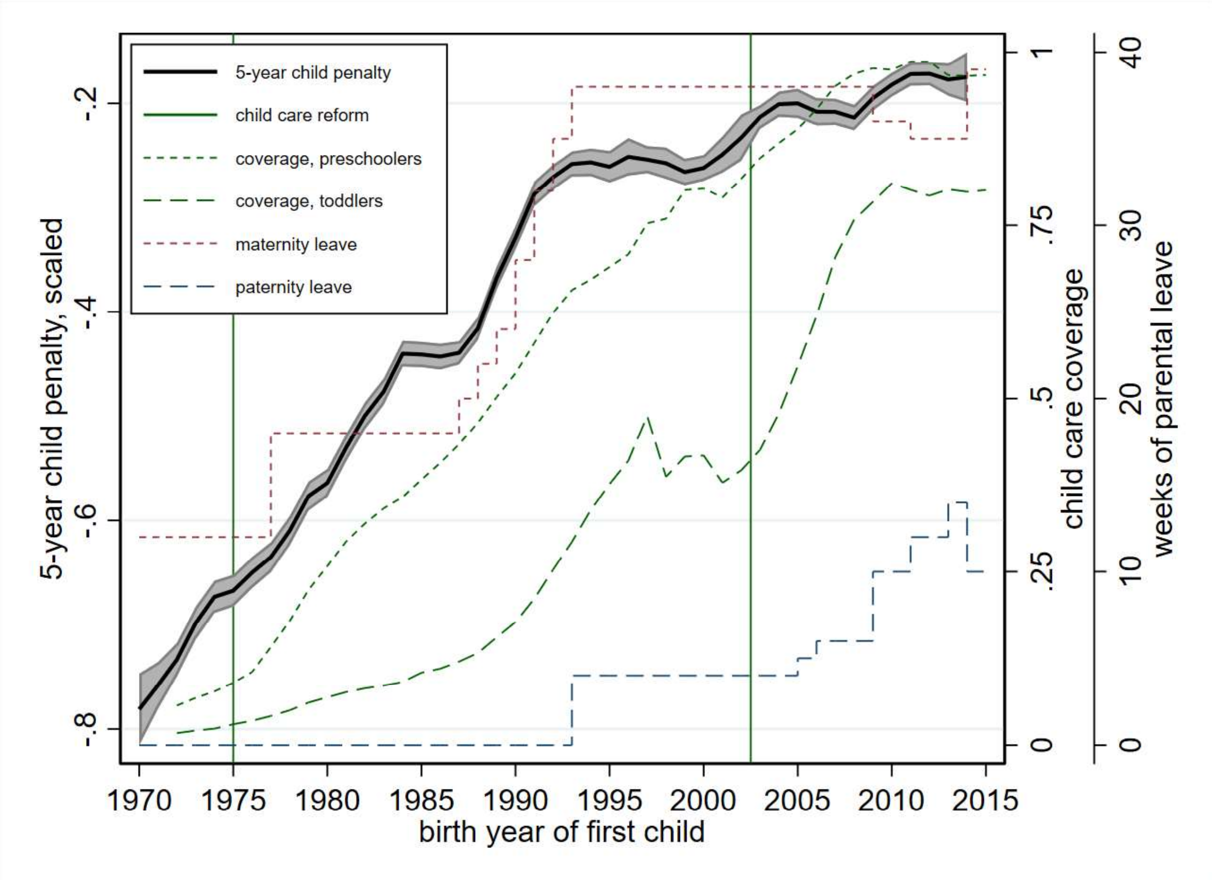
A large literature shows the arrival of children causes sharp drops in earnings for mothers but not fathers (Angelov *et al.*, 2016; Chung *et al.*, 2017; Andresen and Nix, 2022; Kleven *et al.*, 2019a,b). Cortés and Pan (2020) discuss how this well documented disparity in the impacts of children, known as the "child penalty", accounts for a large portion of the remaining gender gap across a variety of countries. As a result, if societies wish to reduce gender income gaps, the first priority is to reduce child penalties. However, can government policies decrease child penalties?

Figure 1 shows our estimates of the cumulative five-year child penalty over the past 45 years in Norway (see Section 4 for estimation details). The figure provides a much longer time horizon than previously documented in the child penalty literature, capturing how the impact of children on earnings of mothers relative to fathers has evolved over the very long run.

The figure shows that the child penalty, denoted by the black line with 95% confidence intervals in gray, has decreased significantly from 1975 to 2015. Mothers used to experience a drop in average earnings relative to fathers following their first child equivalent to almost 80% of their counterfactual earnings in the 1970's (including mothers dropping out of the labor force), which decreased to approximately 20% by 2010. Compared to point-in-time child penalty estimates, this provides a fuller picture and suggests growing equality between parents, although reductions slowed in the last twenty years.

The Norwegian government adopted a number of programs during this period that may have played a role in this reduction. Figure 1 additionally documents these policies, depicting changes in paternity and maternity leave (blue and red dashed lines, respectively, measured on the right-most axis) and two major childcare reforms in 1975 and 2002 indicated by vertical lines, with resulting changes in the quantity of childcare coverage as a share of all children of either preschool or toddler age indicated by dashed green lines on the first right-hand axis. It is interesting to see how reductions in child penalties correlate with expansions of both maternity leave and childcare increases following the 1975 reform. The child penalty is then flat from 1995-200, followed by modest but still relevant reductions in child penalties after 2000 that correlate with paternity leave expansions and childcare expansions for toddlers. Of course, none of these correlations can be interpreted causally.

Figure 1: The Child Penalty Over Time



Notes: Child penalties estimated separately in five year birth cohorts centered on cohorts as indicated by the horizontal axis. 95% confidence intervals in gray. Child penalties are estimated using the event study framework from equation 1, and measure the sum of the impacts of a child's birth on mother's income relative to the father's income from the year of birth of the first child until the year the child turns five years, compared to the difference between mother's versus father's income the year before the child's birth, scaled with the predicted earnings of the mother absent children over the same period. For more information, see Section 4.1. Major policy interventions indicated by the colored lines and the figure legend. Paternity leave is measured as the paternity leave quota, while maternity leave is measured as the sum of the maternity leave and shared quotas, which reflect the most typical use of these quotas.

In the remainder of this paper we examine two of the more recent policies depicted for which we have detailed data and research designs to identify their causal impacts on the child penalty: paternity leave and government subsidized childcare. As a means of increasing fathers' involvement in raising children, the "daddy quotas" (leave reserved for fathers) of the Scandinavian countries have attracted considerable interest.[1] Similarly, there has been significant discussion of the impact of subsidized childcare on gender gaps. The descriptive results in Figure 1 suggest reductions in the child penalty corresponding with the 2002 childcare reform (second vertical green line), but it is not clear if this impact is causal. Similarly, the later paternity leave expansions appear to coincide with small reductions in the child penalty, but these could be spurious correlations. Understanding whether these policies caused reductions in the child penalty is the main focus of this paper. Note that we do not examine causal impacts of maternity leave for three reasons.[2] First, leave is not observed for earlier years, so we could only estimate reduced form impacts. Second, external validity of changes in the 1970s and 1980s to other countries today is questionable, making such an exercise less pertinent. Third, maternity leave has already been adopted by almost all countries (the United States notably excepted), so evidence on paternity leave and childcare expansions is more relevant as other countries consider their adoption.

There are theoretical reasons why these two policies could reduce child penalties. Paternity leave could increase utility fathers get from time with children and change gender norms around childcare. Subsidized childcare reduces the cost of market care in the household budget constraint and could increase labor supply of mothers if they are the counterfactual caretaker. Whether these policies reduce child penalties in practice is an empirical question. We use administrative data from Norway to estimate causal impacts of each policy in the same setting, allowing us to compare and contrast how each impacts the child penalty.

To identify the impact of paternity leave on the child penalty we use a robust stacked regression discontinuity design to estimate the impact of six reforms to the paid paternity leave quota in Norway from 2005 to 2014, illustrated in blue dashed lines in Figure 1. We estimate a strong first stage: the reforms significantly increase paternity leave takeup by 0.85 week per week of quota. Despite this, we find no significant impact on the child penalty. Point estimates are approximately zero and we can rule out effect sizes larger than around 5% reductions in child penalties per week of paternity leave. We also find no impact of paternity leave on norms around childcare

[1] A few other countries introduced similar quotas, including Ireland (14 weeks), Slovenia and Iceland (13 weeks), Germany (8 weeks), Finland (7 weeks), and Portugal (6 weeks) (OECD (2014)).

[2] These maternity leave expansions were explored in Dahl *et al.* (2016); Carneiro *et al.* (2015), but not impacts on child penalties.

within the couple, as measured by fathers' takeup of parental leave for subsequent children. We can rule out effects larger than 0.1 week per week of leave taken for the initial child. Combined, these results suggest paternity leave has limited potential to reduce child penalties and does not substantially change norms within couples.

To identify the impact of access to high quality childcare on child penalties we use a large-scale Norwegian reform from 2002 that expanded childcare availability for 1- and 2-year-olds, as illustrated in Figure 1. The reform increased subsidies to childcare institutions, leading to a rapid expansion of previously rationed care slots. We exploit the variation across municipalities and over time in construction of new slots and centers, instrumenting individual childcare use with the rationed, municipality-level availability of slots in a variation of the setup in Andresen and Havnes (2019). Unlike Andresen and Havnes (2019), we estimate impacts on the child penalty directly and do so for first time parents, who are most relevant for the child penalty. Our results indicate that formal care reduces the child penalty by 23% during the years of early childcare use. Although impacts are not persistent in the long run, the sizable short run effects suggest that subsidized childcare could play an important role in reducing the child penalty and the gender income gap.

Our paper is most closely related to the literature on child penalties. Large child penalties are found in Norway (Andresen and Nix, 2022) , the United States (Chung *et al.*, 2017), Sweden (Angelov *et al.*, 2016), Denmark (Kleven *et al.*, 2019b), Chile (Berniell *et al.*, 2021), Mexico (Aguilar-Gomez *et al.*, 2019), Austria, Germany, and the United Kingdom (Kleven *et al.*, 2019a). As these papers demonstrate, the child penalty is pervasive, occurring in every country in which it has been studied, and is a pivotal component of the gender income gap (Cortés and Pan, 2020). In this paper we provide evidence on whether government policy can reduce the child penalty.

We also contribute to a large literature on the impact of paternity leave and subsidized early childcare on a range of outcomes. Olivetti and Petrongolo (2017) provide a good review of the current literature on the economic consequences of family leave policies. Most directly, we contribute to the literatures on the effects of child care and paternity leave policies on parents' labor supply. The literature that examines the impacts of leave on parents' earnings and labor supply typically finds mixed results (for an overview, see Rossin-Slater, 2017).[3] Regarding childcare, a large literature (summarized in Blau and Currie, 2006; Akgunduz and Plantenga, 2018; Morrissey, 2016) contains a range of estimates on the elasticity of female labor supply to childcare availability. Of most relevance here are Havnes and Mogstad (2011) who find small effects from a childcare reform for preschoolers, and Andresen and Havnes (2019) who find considerably larger effects from the same reform for toddlers that we exploit in this paper.

We contributeby isolating impacts of these policies on child penalties directly and estimating impacts of both policies in a single setting. While the prior literature largely focus on maternal (and less frequently paternal) labor supply in isolation, neither positive effects on mothers' earnings nor negative effects on fathers' earnings is sufficient to establish reductions in child penalties, which are fundamentally about relative earnings. Thus, positive impacts on mother's earnings might be counteracted by positive impacts on father's earnings, resulting in no changes in child penalties. Conversely, as we find for formal care, significant impacts on child penalties may be missed because the impact on each parent's earnings separately is insignificant. We estimate child penalties and impacts of policies across spouses and event times jointly to account for correlations of these parameters and establish impacts on the child penalty itself.

2 Parental Leave and Childcare in Norway

Norwegian parents have been entitled to paid parental leave since the 1930s, with expansions to 12 weeks in 1956 and 18 weeks in 1977. Total parental leave is currently 49 weeks at 100 percent replacement or 59 weeks at 80 percent replacement rate. The length of leave steadily increased since the mid 1980s, as Figure 1 illustrates, with paternity leave given by the blue dashed line. Appendix Table A1 reports details on every leave reform in Norway, 1992-2014. The maternal

[3] Rege and Solli (2013) find a decrease in fathers' earnings long term in Norway from a 1993 reform using a DiD approach, Druedahl *et al.* (2019) show a Danish increase in the daddy quota increased mothers' share of household earnings; Johansson (2010) finds a Swedish policy increased mother's earnings (no impact on fathers); Ekberg *et al.* (2013) find fathers are no more likely to take sick leave using a Swedish reform; Cools *et al.* (2015) estimate effects of paternity leave extensions in Norway and, like us, find no effect on traditional labor supply allocations within family although children's test scores improve; Andersen (2018) finds that father's leave reduces the within household gender gap in Denmark. Patnaik (2019) finds a large change in the division of household labor from a Canadian paternity leave expansion. Bana *et al.* (2018) find that men take much less paid family leave than women in California

and paternal quota columns report parental leave in weeks reserved exclusively for the mother and father. Remaining leave in column 5 can be shared among parents however they choose, but most commonly taken by mothers, implying total possibly maternity leave allowed in column 6 and depicted in Figure 1 by the red dashed line. Leave benefits are capped at 640,000 Norwegian kroner (NOK)[4], roughly 75,000 USD, with many employers topping up.

To qualify for leave, a parent must have been employed for minimum 6 of 10 months prior to birth, and annual earnings must exceed a low threshold of around 50,000 NOK or 6,000 USD. For fathers to take leave, both parents must qualify. Mothers who do not qualify for parental leave receive a one-time-benefit of 90,300 NOK (approximately 10,000 USD). In addition to paid leave, all parents have job protection for another year of unpaid leave.

Following parental leave, Norway has a well-regulated and highly subsidized childcare sector. Because of heavy subsidies for formal care, the market for paid childcare outside this system is very small, but subsidies are available for both private and public childcare. For children not in formal care and aged 13 to 35/36 months there is a monthly cash-for-care benefit.[5] Figure 1 depicts childcare coverage rates over time in Norway, separately for toddlers (age 1-2) and preschoolers (age 3-5). The formal care sector for preschoolers was well developed in Norway by the early 2000's, with over 80 percent of Norwegian 3-5 year olds attending care, partly due to the 1975 reform depicted. For toddlers, however, coverage was much lower at around 30 percent, and the market was strongly rationed.

The underrepresentation of children below 3 in formal care was the impetus for the Childcare Concord in 2002, a broad, bipartisan agreement to increase availability of toddler care. Following this reform, coverage increased rapidly for 1- and 2-year-olds over the next years as shown in Figure 1. This expansion varied considerably between municipalities and over time(see Appendix Figure B2), making the expansion of care availability a potential instrument for the endogenous choice of how much childcare to use. We use this for identification.

[4]All monetary values are reported in 2020 NOK, adjusted using the CPI.
[5]2001-2009 the benefit was around 3,500 NOK, equivalent to ~550 USD.

3 Data and Sample

We use data from Norwegian administrative registers covering the entire resident population for 1967-2019. Our main outcome of interest is the child penalty, which we estimate from annual labor market earnings (including taxable benefits) of fathers and mothers. For long-run estimates in Figure 1 we use pensionable income, a slightly broader income measure. We also use data from the Norwegian Public Insurance system (FD-Trygd) which contains information on all leave spells.[6] The data does not contain direct links to the child for whom leave is taken, only to the individual who takes leave. We therefore infer the relevant child from birthdates, which is straightforward given we limit attention to first born singletons. We assign a parent's leave spell to a particular child if it starts between 60 days before or 3 years after the child is born, and 60 days before any subsequent child is born. This mirrors rules for parental leave, which can be taken up to the age of three, but leave not taken by the next child's birth is lost. We match 99.45 percent of all leave spells to a child. We treat 80 percent and 100 percent compensated leaves identically, as they are not distinguishable in the data and both reflect time away from work.[7]

To construct our sample, we take all children born in Norway in 2001 to 2006, 2009, 2011, 2013 and 2014, the relevant years for our policy reforms. We restrict to children where we know this is the first child and the identity of both parents. We exclude a small number of children with same-sex parents, and include only children whose parents both resided in Norway the year before birth. To precisely determine our two treatment variables, we also restrict attention to singleton children. To measure labor supply before and after birth we exclude parents who are not 22-60 at the time of birth. We do not restrict the sample based on number of siblings, as we view future fertility as a potentially endogenous outcome of the reform. For the parental leave analysis, we only include couples where the mother took some regular leave, so that the father may also qualify.[8] This leaves us with a sample of around 92,000 couples for the childcare

[6]There are five types of leave spells recorded. More than 97 percent of the leave spells are regular leave spells, on which we focus.

[7]In cases of graded leave (working part-time) we compute efficient days at home. We cap the small number of parents who appear to be taking longer than the total leave (1.15 percent of mothers and 0.08 percent of fathers) to 60 weeks.

[8]At first glance, this might seem like an endogenous sample selection, but this is imposed to determine eligibility.

analysis and almost 90,000 couples for the paternity leave analysis. To measure our instruments in the formal care application, we use the municipality of residence of the child at the beginning of the year the child will eventually turn 2.

As Section 4 describes, the paternity leave sample is further refined using optimal bandwidths around the timing of the reform. We expand these samples to a panel from 4 years before birth to 5 years after birth, keeping parents in the sample only when they reside in Norway. We match these parent-years to labor earnings from tax records, measuring these in 2020 NOK using the CPI. Finally, we set a handful of negative labor earnings to zero and winsorize around 0.25 percent of the sample to NOK 2 million to restrict a very small number of extreme earnings that otherwise distort estimates. For the paternity norms applications, we construct a sample of fathers of all children (not just first births) born in the reform years except 2014, which is excluded because we need to measure paternity leave use for future children. We restrict attention to fathers who have another child (and not more than one) in 2015 or earlier, so that we can measure paternity leave use for the next child. This leaves 53,000 fathers.

Finally, to estimate long run child penalties (Figure 1), we construct samples of first births from 1970-2014 following the procedure described above. Unfortunately, we need to rely on pensionnable income to measure earnings in earlier years, a broader income measure. Furthermore, we cannot measure residency and earnings data does not include individuals with zero earnings. Thus, we include children where both parents are born in Norway or where both parents immigrated to Norway more than four years before birth and emigrated from Norway more than five years after birth. We set missing earnings to zero. This leaves approximately 577,000 children.

4 Identifying Child Penalties and Policy Impacts

4.1 Identifying Child Penalties

To identify child penalties, we estimate a standard event study model:

There are no incentives not to take up maternity leave if you qualify, because there are virtually no alternative care options available in the first six months so a parent must stay home and maternity leave far exceeds the alternative one-time benefit ineligible mothers receive.

$$y_{it} = \overbrace{\sum_{j \neq -1} \beta_j \mathbb{1}[t - E_i = j]}^{\text{Earnings dynamics for fathers}} + \overbrace{\sum_{j \neq -1} \alpha_j \mathbb{1}[t - E_i = j, m(i) = 1]}^{\text{Child penalty}} + \theta_{m(i)a(it)} + \gamma_{m(i)t} + \epsilon_{it}, \quad (1)$$

where $m(i)$ is a dummy equal to 1 for mothers (0 for fathers). $\mathbb{1}[A]$ is the indicator function for event A. α_j captures the child penalty (the difference in the impact of children on mothers versus fathers). β_j captures earnings dynamics around birth for fathers and $\alpha_j + \beta_j$ captures earnings dynamics for mothers. Andresen and Nix (2022) discuss how the assumptions required for β_j and $\alpha_j + \beta_j$ to be interpreted causally are problematic given pervasive pre-trends in event studies looking at impacts of children on mothers' and fathers' earnings separately. In contrast, to interpret α_j causally, we require a standard common trends assumption: in the absence of children, earnings of mothers and fathers would have followed the same trend conditional on age profiles and yearly shocks. We also control flexibly for yearly shocks, γ_{mt}, that vary freely for mothers and fathers and flexible age profiles, θ_{ma}, accounting for the fact that most couples tend to have their first child when earnings are increasing. Standard errors are clustered by couple and robust to heteroskedasticity.

As a summary measure of the child penalty in Figure 1 we calculate the 5-year child penalty, which is the sum of the α_j parameters for nonnegative event times. Following Kleven *et al.* (2019b), we scale this with the sum of the predicted earnings for women over the same time period, excluding event time coefficients.

We examine how policies reduce child penalties. This happens either if the policy increases mothers earnings, reduces fathers earnings or, in principle, by affecting any parent's earnings in the base year. Formally, for any causal estimate ϕ_{jm} of the impact of a policy on earnings in relative time j for parent $m \in 0, 1$, the impact of the policy on the child penalty is:

$$\Delta CP_j = \frac{\phi_{j1} - \phi_{j,-1} - \phi_{j0} + \phi_{j,-1}}{\alpha_j}, \quad (2)$$

a measure of the relative change in the child penalty at time j from a one unit increase in the policy of interest. To allow impacts of the policy to be correlated over time and across spouses as well as with the child penalty itself, we stack the specifications for the child penalty and policies described below, estimating them jointly to account for the dependence of the estimators, before constructing our measure of the impact on the child penalty in 2.[9]

4.2 Identifying the Effects of Paternity Leave Using Regression Discontinuities

We use a regression discontinuity design to identify impacts of paternity leave. For each reform year and relative time, and separately for mothers and fathers, we estimate the fuzzy RD model:[10]

$$\begin{aligned} y_i &= \alpha + \phi_L L_i + f_0(x_i)\mathbb{1}(x_i < 0) + f_1(x_i)\mathbb{1}(x_i \geq 0) + \mu_i \\ L_i &= \pi + \gamma Q_i + g_0(x_i)\mathbb{1}(x_i < 0) + g_1(x_i)\mathbb{1}(x_i \geq 0) + \xi_i \end{aligned} \tag{3}$$

where x_i, the running variable, is the number of days after the reform date that the child was born. $f(x_i)$ and $g(x_i)$ are local linear polynomials that are separate on either side of the cutoff. Q_i is the paternity leave quota, which scales first stage estimates to be comparable across reforms, and only varies at the cutoff within a reform year. We use the optimal bandwidth that minimizes the mean squared error of the RD estimate to define the sample of births we use, and a triangular weighting function to obtain estimates local to the cutoff (Calonico *et al.*, 2014). When conducting bias-correct inference to account for the fact that the f and g functions are approximations to the unknown relationships between the outcome and the running variable, we use local quadratic polynomials and alternative bandwidths (Calonico *et al.*, 2014, 2018).

Our identification strategy exploits the fact that parents of children born just before the reforms were not subject to changes in parental leave quotas, whereas parents of children born right after each reform were subject to the changes. The reforms were announced in October the

[9] The first reference we have seen of this approach to accounting for the covariance between estimators is Erickson and Whited (2002).

[10] In practice, we estimate all relative time- and parent-specific estimates in a stacked regression, clustering at the birthdate of the child to allow arbitrary correlation of errors within couple and over time as well as on the running variable.

year before implementation as part of the budgeting process, making it nearly impossible to plan conception in response to announced quota changes in order to manipulate birth dates around the cutoff in July. Appendix Figure A2 verifies there is no statistically significant change in the density of births around the cutoff. The critical assumption for the validity of our RD approach is that underlying regression functions are continuous at the threshold. This implies that the population of couples around the discontinuity are identical. Balancing tests in Appendix Table A2 provide empirical support for this assumption.[11]

To increase precision, our main results combine the six reforms. The common way of doing this is to re-center the running variable to be zero at the relevant cutoff for all individuals and run semiparametric RD estimates in the pooled sample. This restricts the functional form of the polynomials and the optimal bandwidth to be the same for each reform, and may not adequately control for the underlying effects of the running variable on the outcome. Alternatively, we can take a weighted average of all reform-specific estimates, using the relative number of observations used at each cutoff as weights (Cattaneo *et al.*, 2016).

A third and more straightforward way to stack estimates is to allow the local polynomials of the running variable to vary by cutoff and use cutoff-specific optimal bandwidths from the individual specifications.[12] This produces an inverse variance weighted average of the reform specific fuzzy RD-estimates, allowing us to put more weight on more precise RD estimates. This is particularly important in our setting because the size of the reform varies considerably, with the quota increasing by 1, 2 or 4 weeks and even decreasing by 4 weeks. As expected, larger reforms in 2009 and 2014 provide more power, and thus are given more weight in stacked estimates, (see Appendix Table ??). We report all three ways of combining separate reforms. We prefer the stacked specification because it is more robust than the pooled specification and more precise than the weighted specification. We cannot calculate bias-corrected standard errors for this specifica-

[11]This table reveals an imbalance inmaternity leave take-up. Some reforms increase paternity leave quotas at the expense of shared leave taken in practice by mothers. We do not believe these small changes in maternity leave take-up from already high amounts drive our results. In Appendix A.4, we exploit the fact that some reforms expanded paternity leave at the expense of maternity leave, while others expanded total leave length. This allows us to instrument for both maternity and paternity leave use, confirming our baseline paternity leave results.

[12]To avoid weak instruments issues, we use a single instrument measuring the dummy in the stacked specification, but maintain separate local linear controls and bandwidths specific to each reform. Results are very similar if we instead use the overidentified model with 6 reform-specific instruments for a single treatment.

tion, but the problem should be relatively minor because differences between cluster-robust and bias-corrected standard errors are small.

Finally, we must emphasize the local nature of our estimates. They reflect local average treatment effects for compliers: families who take up more paternity leave because of the increased quota. Because fathers were free to take up shared leave, these are *unwilling* users of paternity leave, in contrast to the constrained compliers to the childcare expansion. However, we emphasize that for both applications, we view compliers as the policy-relevant group.

4.3 Identifying the Effects of Early Childcare Using Instrumental Variables

To identify the impact of early childcare we exploit the sizable 2002 reform in Figure 1 and the fact that the expansions of care facilities varied considerable over time and space in the years following the reform for reasons that seem to be uncorrelated with underlying drivers of labor supply (Andresen and Havnes, 2019). The variation in expansion across time and space is illustrated in Appendix Figure B2. Specifically, we instrument for the actual use of care over the toddler period by using measures of the availability of slots in care, controlling for municipality and year fixed effects. Separately for mothers and fathers and each event time, we estimate the following IV model:[13]

$$
\begin{aligned}
y_i &= \pi_{k(i)} + \gamma_{t(i)} + \beta_{a(i)} + \theta_{b(i)} + \phi_c m_i + \epsilon_i \\
m_i &= \tilde{\pi}_{k(i)} + \tilde{\gamma}_{t(i)} + \tilde{\beta}_{a(i)} + \tilde{\theta}_{b(i)} + \gamma_1 CC^1_{k(i)} + \gamma_2 CC^2_{k(i)} + \tilde{\epsilon}_i
\end{aligned}
\tag{4}
$$

where γ_t are calendar year fixed effects, π_k are municipality fixed effects for the municipality of residence at the time of treatment and m_i is our measure of childcare use in years from ages 13 to 36 months. β_a and θ_b are age fixed effects for the parent (in years) and calendar birth month effects for the child, which increases precision of our estimates slightly but has little impact on

[13]In practice, we estimate all relative time- and parent-specific estimates in a stacked regression, clustering at the municipality to allow arbitrary correlation of errors within couple, over time, and within cells where the instruments are determined.

point estimates. The instruments are CC_k^1, the share of slots for toddlers in the municipality at age 1 to the population of toddlers, and CC_k^2, the same share measure at age 2.

The variation we exploit comes from variation in expansion of care across municipalities and over time. As long as the exact timing of expansion of care is uncorrelated with other drivers of the child penalty, expansion of care only affects labor supply through use of care, and no children opt out of care when access is expanded, our approach recovers causal effects of an extra year of early childcare on the child penalty for the compliers: the families who take up the newly expanded slots. Because childcare was strongly rationed before the reform, it is natural to think of compliers as the mothers of children who wanted childcare before the reform, but were restricted by the low supply. This can be interpreted as willing, but constrained families, in contrast to the compliers to the paternity leave reforms who take up treatment only because they are "forced to" by the quota. Andresen and Havnes (2019) shows the exact timing of expansion was subject to a range of constraints that were hard to predict, and the timing of expansion was not necessarily easy to predict even for the municipalities themselves. Appendix Figure B1 provides support for the idea that expansions did not systematically vary across municipalities with different pre-reform characteristics (except, of course, the initial coverage rate).[14]

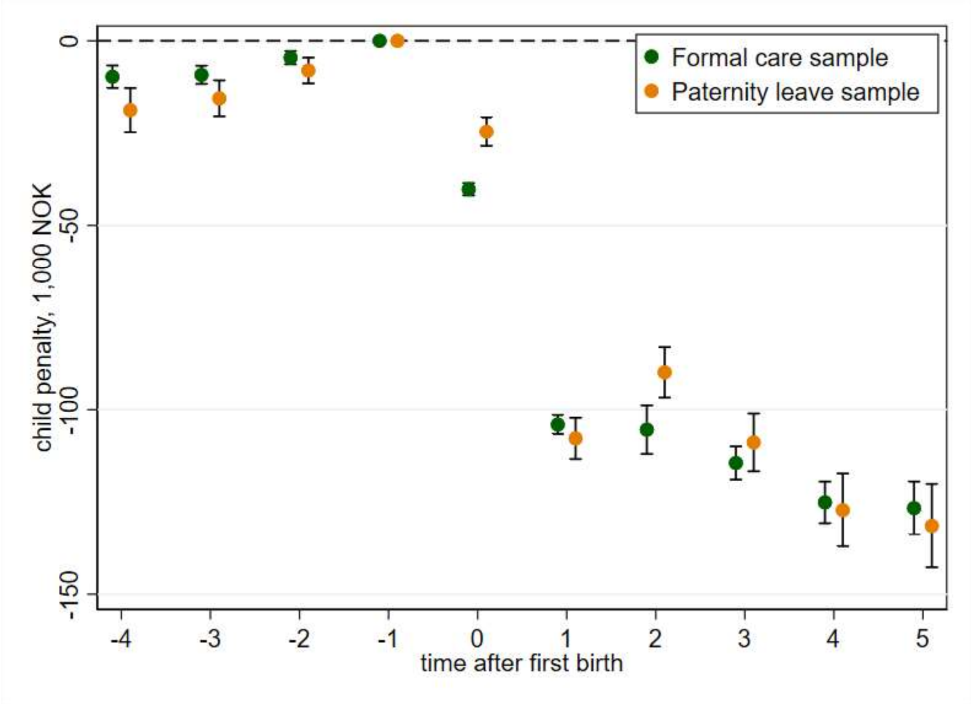
5 Results

In Figure 1 we estimated scaled 5-year child penalties in rolling 5-year intervals for 1970-2015. This shows large reductions over time and covers before, during, and after paternity leave and early childcare policies. In this section we focus on the more recent years where gains have been smaller but still noticeable. We do so both due to data constraints on earlier years and also because the more recent policies are salient for many other countries. In Figure 2 we estimate child penalties for the paternity leave and childcare estimation samples, again estimating α_j from equation 1, but this time reporting estimates separately for four years before and five years after birth. While pre-birth estimates are marginally non-zero, pointing to possible violations of common trends, these are tiny compared to the massive drops in earnings following birth, and thus

[14] Andresen and Havnes (2019) provide additional robustness.

unlikely to be a major concern. The estimates show that mothers' earnings drop over 100,000 NOK more than fathers for all years after birth (relative to the year before birth). We next turn to causal impacts of paternity leave and expanded access to childcare on these child penalties.

Figure 2: Child Penalties for Our Samples



Notes: Figure shows child penalties for the formal care and paternity leave estimation samples using the event study framework from equation 1. These are the data samples that we use to estimate causal impacts of each policy on the child penalty.

Table 1 reports first stage estimates for both policies. Panel A shows clear effects of each and every paternity leave reforms on the take up of paternity leave, with first stage coefficients around 0.9. Our preferred stacked model stacks all reform-specific models, which is equivalent to an inverse-variance-weighted average of the reform-specific IV estimates. As a summary measure, the weighted average of the first stage estimates is 0.91 weeks of additional paternity leave per week of quota. The joint first stage F-statistic is around 37, well above conventional levels[15]

Panel B reports first stage estimates from equation 4 for formal care. Expansions of care at age 1 and at age 2 have strong impacts on early childcare use, with an additional slot in care at age 1

[15]Our first stage also passes the Olea-Pfluger weak IV test, with an efficient F-statistic of 35 compared to a critical value for a maximum 5% bias of 26.1 (Olea and Pflueger, 2013; Pflueger and Wang, 2015).

increasing care use by around 1 year and at age 2 by about 0.6 years. Note that our endogenous variable captures the intensity of use throughout the full period while the instruments are measures of availability at two particular points during this period, and changes in the instruments typically happen in August each year at the start of the school year. The IV strategy thus scales the reduced form estimates to reflect a full year of early childcare use. The F-statistic is around 170, indicating a very strong first stage.

Table 1: First Stage Estimates for Parental Leave and Childcare

Panel A: Weeks of paternity leave

	2005	2006	2009	2011	2013	2014	Pooled	Weighted	Stacked
RD estimate per week	0.965***	1.130***	0.992***	0.870***	0.579**	0.735***	0.85***	0.870***	0.909***
cluster-robust std. err.	(0.330)	(0.343)	(0.0876)	(0.201)	(0.263)	(0.102)	(0.0342)	(0.0956)	(0.0612)
bias-corrected std. error	[0.377]	[0.412]	[0.104]	[0.234]	[0.314]	[0.122]		[0.113]	[0.0738]
F-statistic	8.55	10.8	128.4	18.8	4.83	52.1	615.3	82.8	215.8
-, bias-corrected	7.53	6.76	91.7	17.5	4.96	50.5		61.2	156.4
Observations	13,730	14,244	15,409	15,335	15,340	15,800	89,858	89,858	
Optimal bandwidth	49.50	55.35	72.11	40.53	65.48	47.45	66.74	57.27†	
Efficient observations	3,836	4,589	6,373	3,683	5,970	4,410	34,577	28,861	
weight in stacked	0.02	0.03	0.60	0.07	0.05	0.23			
weight in weighted	0.13	0.16	0.22	0.13	0.21	0.15	Joint F, all reforms		37.3

Panel B: Years of early care

Coverage, age 1	1.037***	(0.067)	Joint F	167.1
Coverage, age 2	0.594***	(0.056)	Observations	92,767

Notes: Panel A shows semiparametric RD estimates of the effect of paternity leave reforms on paternity leave take-up using optimal bandwidths, triangular kernel, and local linear polynomials on either side of the cutoff. All estimates are scaled to reflect one week of quota increase. Pooled estimates are from a separat model that recenter the running variable and control for a single local linear polynomial and birth year dummies. Weighted estimates reflect the weighted average of of the individual reform specific first stage coeficcients, weighting with the relative number of observations used. Stacked estimates refrer to the inverse-variance-weighted average of the reform-specific first stage estimttess. Conventional standard errors are robust to clustering at the running variable and heteroskedasticity, but not bias-corrected. Bias-corrected inference in brackets. †indicates average bandwidth. Panel B shows the first stage estimates for the childcare reform from equation (4). Standard errors in parentheses, clustered at municipality and robust to heteroskedasticity. Both panels report first stage estimates for mothers in the year before birth. Estimates for fathers and other years may differ slightly due to alternative optimal bandwidths and slight variations in sample due to our unbalanced panel, but are always very similar. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, based on cluster robust, but not bias-corrected inference.

Before turning to the child penalty, our main outcome of interest, we briefly discuss impacts on individual earnings. Appendix Figure A3 shows largely insignificant impacts on mother's and fathers earnings pre-birth. This is reassuring: The variation in childcare and paternity leave exploited by our instruments seems to be unrelated to earnings before birth. Following birth, impacts are close to zero for paternity leave for both fathers and mothers. For childcare, however, we find a clear pattern of reduced earnings for fathers and increased earnings for mothers during the years when childcare is used. While these estimates are not separately statistically significant, what matters for the child penalty is the difference between the impact on mothers and fathers, relative to any impact before birth. These results and what follows indicate the importance of examining the child penalty as a seperate economic parameter of interest.

Figure 3 reports impacts of paternity leave on the child penalty, with estimates shown by orange circles. The y-axis represents the percentage change in the child penalty. Point estimates are zero, suggesting no impact of paternity leave on the child penalty. This zero is relatively precise. The lower bound of the confidence intervals rules out reductions larger than around 5% of the child penalty per week of paternity leave at child age 1-5.

Despite not impacting child penalties, paternity leave could influence gender norms around the distribution of childcare. One measure to capture this is increased use of shared leave by fathers for future children. We exploit the fact that many of the fathers that have a child around the time of the paternity leave reforms subsequently go on to have more children. We estimate our fuzzy RD model using as an outcome the father's total leave take-up for the next child for all children born up to and including 2014 in a setup similar to the peer effects estimates from Dahl *et al.* (2014).[16] Table 2 shows no evidence of permanent impacts on norms as measured by take-up of paternity leave for later kids. Focusing on our preferred stacked estimates, results indicate a non-significant effect of .10 *less* weeks of leave for subsequent children for each week of paternal leave taken for the first child. 95% confidence intervals allow us to rule out effects above around 0.1 week extra paternity leave, suggesting paternity leave use has very limited, if any, impact on norms around childcare as measured by later leave taking behavior.

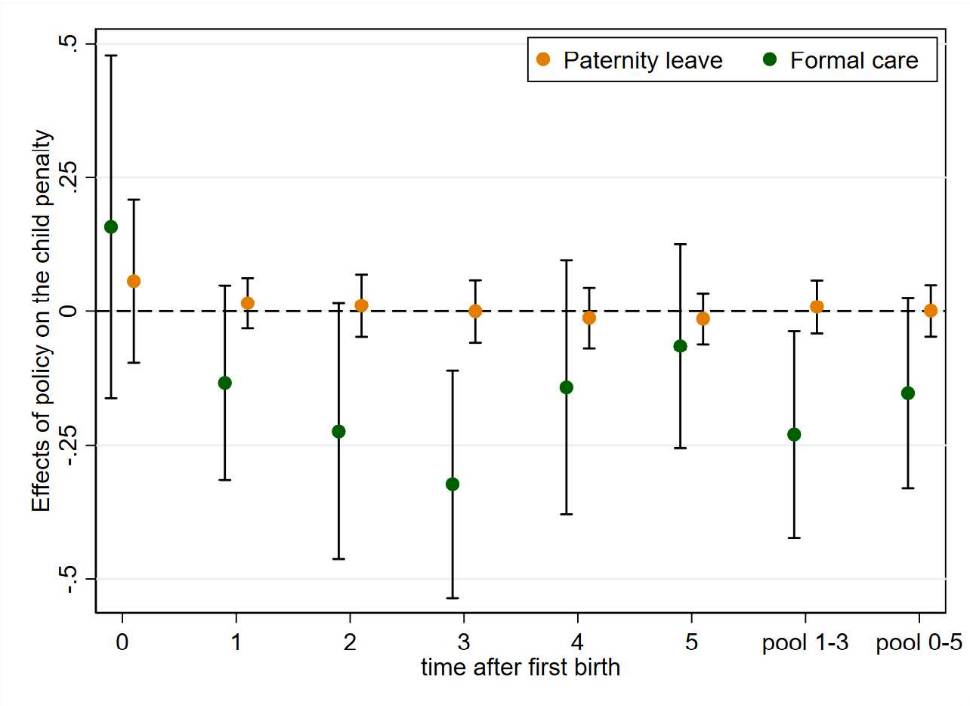
[16]Notice that if fertility was endogenous to the parental leave reforms, this might constitute an endogenous sample selection criteria. Hart *et al.* (2019) investigate fertility response to the 2009 reform and find no evidence of such effects.

Figure 3 also reports estimates of the impacts of a full year of childcare use on the child penalty in green. Results show the child penalty is reduced by around 25-30 percent in years 2 and 3, when the majority of early childcare is used, but the impact appears only in years of treatment and does not persist thereafter. This impact is significant at age 3. For this age, the effect of childcare on the child penalty is also marginally significantly different from the zero impact of paternity leave. Note that lack of persistence after age 3 could be due to all families being treated eventually as they have subsequent children, washing out the impacts. To summarize if we pool the impacts across the years of treatment (age 1-3), we find the child penalty is reduced by 23 percent per year of early care ($p = 0.019$) (the pooled estimate in the figure is for years 1-5 and given the fade-out in years 4-5, it is not significant).

One might argue that the two treatments are not fully comparable, with the impact of childcare in years and paternity leave in weeks, reflecting how these treatments are typically discussed. For a fairer comparison, we could scale paternity leave estimates to have identical public costs as a year of childcare. In 2008, the last year of our childcare sample, municipal and state subsidies for a slot were approximately 160,000 NOK. Subtracting a full year of cash-for-care yields net public costs of early care of 110,000 NOK. The average earnings of fathers in our paternity leave sample implies a public cost of 10,000 NOK for a week of paternity leave. Thus the public cost of 11 weeks of paternity leave equals the public cost of a year of childcare. Considering this, one might scale estimates of paternity leave impacts by 11. After doing this, point estimates are still close to zero, but the 95% confidence interval extends to include reductions in the child penalty around 60%. Nonetheless, we emphasize that there is no evidence paternity leave affects earnings of either mothers or fathers, and point estimates are approximately zero.

We conclude that early childcare shows more promise as a policy tool for reducing child penalties than paternity leave, although it does not have a permanent impact beyond the years of treatment.

Figure 3: Impacts of Paternity Leave and Formal Childcare on the Child Penalty and Norms



Notes: Figure depicts fuzzy RD estimates of the impact of an additional week of paternity leave use on the child penalty using all six reforms in yellow circles. 95% confidence interval, accounting for heteroskedasticity and clustering on the running variable, are indicated in whiskers around the point estimates. Bias-corrected 95% confidence intervals shown in shaded band. Figure also depicts IV results from equation (4) reflecting the impact of an extra year of early childcare use at ages 13-36 months on the child penalty across child age in green circles. Panel (b) shows effects of paternity leave on use of leave for later children, estimated from combining five regression discontinuities.

Table 2: Impacts of Paternity Leave on Leave Taken for Subsequent Children

	Pooled	Weighted	Stacked
RD estimate	-0.044	-0.400	-0.104
cluster-robust std. err.	(0.072)	(0.464)	(0.103)
bias-corrected std. err.		[0.548]	[0.124]
top of 95% CI	0.097	0.510	0.097
–, bias-corrected		[0.575]	[0.107]
Observations	52,896	52,896	
Bandwidth	63.34	60.02	
Efficient N	19,486	18,264	

Notes: Table depicts effects of paternity leave on use of leave for later children, estimated from combining five regression discontinuities as described in the text. 95% confidence interval, accounting for heteroskedasticity and clustering on the running variable, in parentheses, bias-corrected standard errors in brackets.

6 Conclusion

We provided historical evidence of the change in the child penalty over time in Norway and correlations with major reforms to family policies. We find that while fathers take more paternity leave when exposed to a non-transferable quota, paternity leave has no impact on the child penalty or on leave taken for subsequent children, pointing to limited impact on gender norms. While there may exist other benefits of paternity leave (Persson and Rossin-Slater, 2019), if the goal of paternity leave is to reduce the gender income gap, our results suggest it does not achieve this objective.

In contrast, we show early childcare use reduces the child penalty by 23 percent per year of use in the years of treatment. Moreover, Drange and Havnes (2019) find that early childcare in Norway had large positive impacts on the children's outcomes, although this is not universally true in other contexts (see Fort *et al.* (2020) who find negative effects of daycare on child outcomes in Italy). Our results suggest that if policy makers wish to decrease the child penalty, they should focus on providing better outside childcare to families, not on offering paternity leave to fathers.

References

Aguilar-Gomez, S., Arceo-Gomez, E. and De la Cruz Toledo, E. (2019). Inside the black box of child penalties. *Available at SSRN 3497089*.

Akgunduz, Y. E. and Plantenga, J. (2018). Child care prices and maternal employment: A meta-analysis. *Journal of Economic Surveys*, **32** (1), 118–133.

Andersen, S. H. (2018). Paternity leave and the motherhood penalty: New causal evidence. *Journal of Marriage and Family*, **80** (5), 1125–1143.

Andresen, M. E. and Havnes, T. (2019). Child care, parental labor supply and tax revenue. *Labour Economics*, p. 101762.

— and Nix, E. (2022). What causes the child penalty? evidence from same sex and adopting couples. *Journal of Labor Economics*.

Angelov, N., Johansson, P. and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, **34** (3), 545–579.

Bana, S., Bedard, K. and Rossin-Slater, M. (2018). Trends and disparities in leave use under california's paid family leave program: New evidence from administrative data. In *AEA Papers and Proceedings*, vol. 108, pp. 388–91.

Berniell, I., Berniell, L., De la Mata, D., Edo, M. and Marchionni, M. (2021). Gender gaps in labor informality: The motherhood effect. *Journal of Development Economics*, **150**, 102599.

Blau, D. and Currie, J. (2006). *Pre-School, Day Care, and After-School Care: Who's Minding the Kids?*, Elsevier, *Handbook of the Economics of Education*, vol. 2, chap. 20, pp. 1163–1278.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2018). RdrobUSt: Stata module to provide robust data-driven inference in the regression-discontinuity design.

—, — and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, **82** (6), 2295–2326.

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

Cattaneo, M. D., Jansson, M. and Ma, X. (2017). Simple local polynomial density estimators.

—, — and — (2018). Manipulation Testing Based on Density Discontinuity. *Stata Journal*, **18** (1), 234–261.

—, Keele, L., Titiunik, R. and Vazquez-Bare, G. (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics*, **78** (4), 1229–1248.

Chung, Y., Downs, B., Sandler, D. H. and Sienkiewicz, R. (2017). The Parental Gender Earnings Gap in the United States.

Cools, S., Fiva, J. H. and Kirkebøen, L. J. (2015). Causal effects of paternity leave on children and parents. *The Scandinavian Journal of Economics*, **117** (3), 801–828.

Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. *National Bureau of Economic Research.*

Dahl, G. B., Løken, K. V. and Mogstad, M. (2014). Peer effects in program participation. *American Economic Review*, **104** (7), 2049–74.

—, —, Mogstad, M. and Salvanes, K. V. (2016). What is the case for paid maternity leave? *Review of Economics and Statistics*, **98** (4), 655–670.

Drange, N. and Havnes, T. (2019). Early childcare and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics*, **37** (2), 581–620.

Druedahl, J., Ejrnæs, M. and Jørgensen, T. H. (2019). Earmarked paternity leave and the relative income within couples. *Economics Letters*, **180**, 85 – 88.

Ekberg, J., Eriksson, R. and Friebel, G. (2013). Parental leave: A policy evaluation of the Swedish daddy month reform. *Journal of Public Economics*, **97**, 131 – 143.

Erickson, T. and Whited, T. M. (2002). Two-step gmm estimation of the errors-in-variables model using high-order moments. *Econometric Theory*, **18** (3), 776–799.

Fort, M., Ichino, A. and Zanella, G. (2020). Cognitive and noncognitive costs of day care at age 0–2 for children in advantaged families. *Journal of Political Economy*, **128** (1), 158–205.

Hart, R. K., Andersen, S. N. and Drange, N. (2019). Effects of extended paternity leave on union stability and fertility. No. 899.

Havnes, T. and Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics*, **95** (11–12), 1455 – 1465, Special Issue: International Seminar for Public Economics on Normative Tax Theory.

Johansson, E. (2010). The effect of own and spousal parental leave on earnings. (4).

Kleven, H., Landais, C., Posch, J., Steinhauer, A. and Zweimuller, J. (2019a). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*.

—, — and Søgaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*.

Morrissey, T. W. (2016). Child care and parent labor force participation: A review of the research literature. *Review of Economics of the Household*, pp. 1–24.

NOU 2017:6 (2017). Offentlig støtte til barnefamiliene.

OECD (2014). OECD family database: PF2.1 - key character of parental leave systems. available at http://www.oecd.org/els/soc/PF2_1_Parental_leave_systems_1May2014.pdf.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, **31** (3), 358–369.

Olivetti, C. and Petrongolo, B. (2017). The economic consequences of family policies: lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, **31** (1), 205–30.

Patnaik, A. (2019). Reserving time for daddy: The consequences of fathers' quotas. *Journal of Labor Economics*, **0**, null, forthcoming.

Persson, P. and Rossin-Slater, M. (2019). When dad can stay home: Fathers' workplace flexibility and maternal health. *NBER Working Paper No. 25902*.

Pflueger, C. E. and Wang, S. (2015). A robust test for weak instruments in stata. *Stata Journal*, **15** (1), 216–225(10).

Rege, M. and Solli, I. F. (2013). The impact of paternity leave on fathers' future earnings. *Demography*, **50** (6), 2255–2277.

Rossin-Slater, M. (2017). Maternity and family leave policy.

Online Appendix

A Paternity leave: Robustness and Additional Results

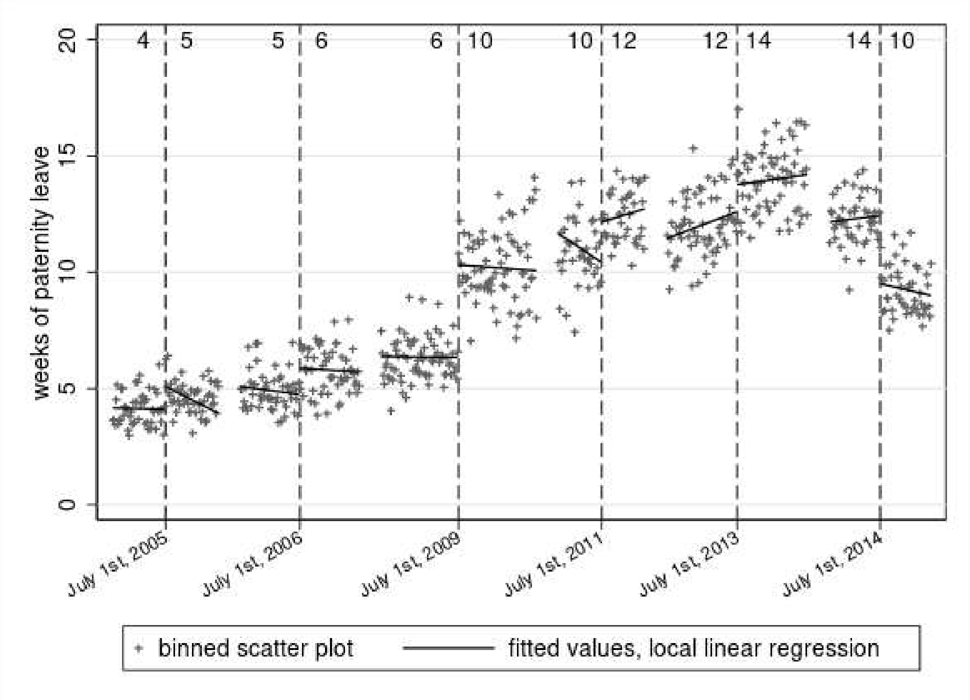
Table A1: Parental Leave Reforms in Norway, in Weeks

(1)	(2)	(3)	(4)	(5)	(6)
Reform Date	Total leave	Maternal quota	Paternal quota	Shared leave	Max leave for mother
April 1, 1992	35 (44.3)	8 (2 before birth)	0	27 (36.3)	35 (44.3)
April 1, 1993	42 (52)	9 (3 before birth)	4	29 (39)	38 (48)
July 1, 2005	43 (53)	9 (3 before birth)	5	29 (39)	38 (48)
July 1, 2006	44 (54)	9 (3 before birth)	6	29 (39)	38 (48)
July 1, 2009	46 (56)	9 (3 before birth)	10	27 (37)	36 (46)
July 1, 2011	47 (57)	9 (3 before birth)	12	26 (36)	35 (45)
July 1, 2013	49 (59)	17 (3 before birth)	14	18 (28)	35 (45)
July 1, 2014	49 (59)	13 (3 before birth)	10	26 (36)	39 (49)

Notes: Parental leave in weeks. Numbers in parenthesis (except maternal quota) indicate weeks of leave if taken at 80 percent compensation, otherwise at 100 percent. *Source:* NOU 2017:6 (2017).

A.1 First Stage Graphically

Figure A1: Fuzzy RD First Stage Estimate



Notes: First stage estimates for each reform, using local linear polynomials, triangular weights and optimal bandwidths. Top numbers are weeks of paternity leave quota around the discontinuity.

A.2 Balancing Tests

Table A2 provides sharp RD balancing tests for a range of covariates using our stacked fuzzy RD model. We re-estimate the optimal weights for each balancing variable in order for the specification to be as similar as possible to our baseline model. All balancing variables are measured the year before birth. Variables are mostly balanced across the cutoff, with the exception of maternity leave use, which we cover below, and which reflect that some of the reforms increased paternity leave quotas at the expense of shared leave, while others increased leave altogether.

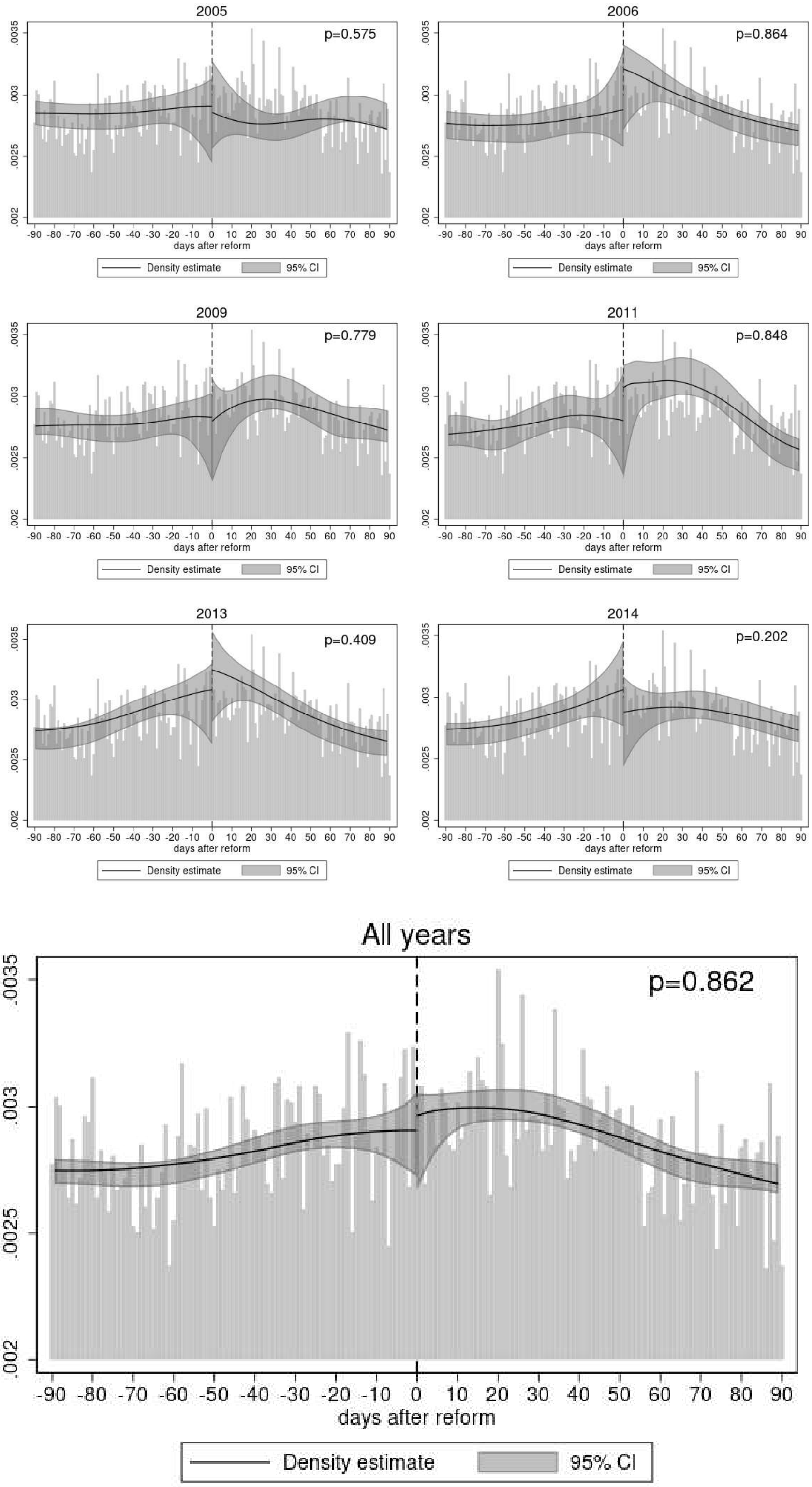
Figure A2 provide robust local polynomial estimates of the density of births around the cutoff.

Table A2: Stacked Fuzzy RD Balancing Tests

	2005	2006	2009	2011	2013	2014	Stacked
Father's age	0.382	0.0328	0.000742	-0.257	0.408	0.0443	0.0243
	(0.336)	(0.264)	(0.0699)	(0.187)	(0.339)	(0.0980)	(0.0518)
Father's years of education	-0.0162	-0.299	0.0561	-0.00160	0.0799	0.124**	0.0548*
	(0.240)	(0.196)	(0.0426)	(0.146)	(0.128)	(0.0536)	(0.0316)
Father's ed. missing	-0.00236	1.84e-14	-0.0000619	0.000416	0.000985	0.000329	0.0000924
	(0.00197)	(2.46e-14)	(0.0000627)	(0.000338)	(0.000864)	(0.000336)	(0.000112)
Mother's age	0.00106	-0.216	0.0140	-0.0901	0.156	-0.0169	-0.00693
	(0.247)	(0.232)	(0.0509)	(0.158)	(0.238)	(0.0873)	(0.0419)
Weeks of maternity leave	0.633	0.0919	-0.374***	-0.116	-0.201	-1.073***	-0.513***
	(0.551)	(0.299)	(0.0981)	(0.221)	(0.375)	(0.147)	(0.0739)
Mother's years of education	0.151	-0.244	0.00583	0.00701	-0.0407	0.0436	0.00561
	(0.179)	(0.198)	(0.0400)	(0.128)	(0.144)	(0.0621)	(0.0317)
Mother's ed. missing	-1.12e-14	0.00163	-0.000519	0.000507	1.40e-13	0.000181	-0.000166
	(0.000)	(0.00177)	(0.000490)	(0.000409)	(0.000)	(0.000183)	(0.000321)

Notes: Each cell contains a separate fuzzy RD estimate of the effect of paternity leave use on balancing variables using optimal bandwidths, triangular kernel and local linear polynomials on either side of the cutoff. Stacked estimates refer to the weighted average of reform reform-specific estimates as discussed in the text. Balancing variables (except maternity leave) are measured the year before birth. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$, based on the cluster robust, not bias-corrected inference.

Figure A2: Density Plots Below and Above Cutoffs

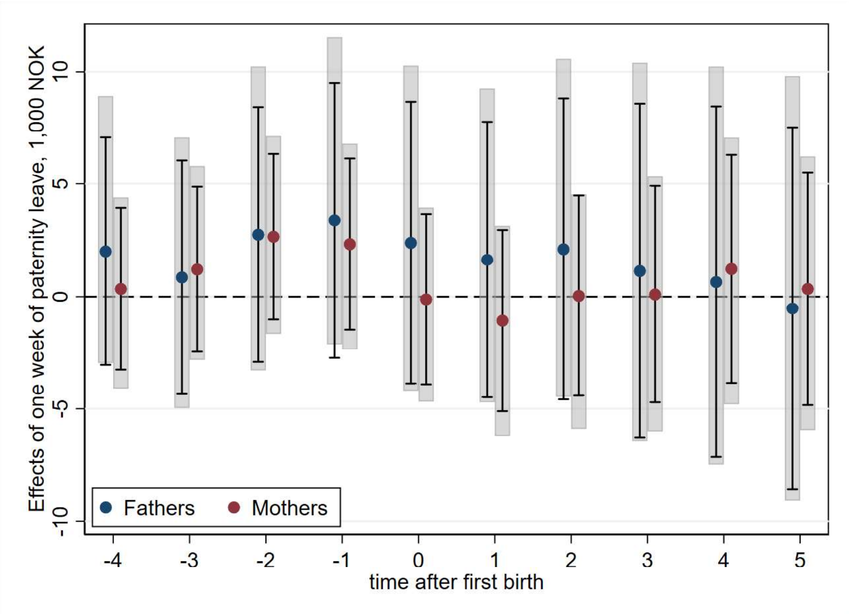


Note: Graphs show density estimates above and below the cutoff using methods described in Cattaneo *et al.* (2017) and implemented in Cattaneo *et al.* (2018). p-values reported are for a bias-corrected test of whether the densities at the cutoffs are equal. Triangular weights and MSE-optimal bandwidths.

A.3 Effects on Earnings Per Year and Parent Type

Figure A3 reports the impacts on mothers' and fathers' annual incomes over time. There is no effect of paternity leave use on pre-birth outcomes. This is a reassuring, and can be interpreted as an additional placebo test. Following birth, we see no impact of paternity leave use at years 0 and 1 on the labor income of mothers or fathers when most of the leave take-up happens. Nor do we see any impact in the following years; the estimates are flat and centered at zero. Using the stacked specification and conventional, cluster-robust inference we can rule out positive impacts larger than around NOK 5,000-8,000 on mother's annual earnings in response to each week of paternity leave use for all years post-birth.

Figure A3: RD Estimates of the Effects of Paternity Leave Use on Mothers' and Fathers' Earnings.



Notes: Red circles indicate stacked fuzzy RD estimates of the impact of an additional week of paternity leave use on mother's earnings over time. Blue circles show the same for fathers. 95% confidence interval, accounting for heteroskedasticity and clustering on the running variable indicated in whiskers around the point estimates. Bias-corrected 95% confidence intervals shown in shaded bands.

A.4 Accounting for Effects of Maternal Leave

As evident from Table A1, several of the reforms affected not only the paternity leave quota, but also the maternity leave quota and the shared leave. As documented in Table A2, this resulted in reduced maternity leave take-up at the cutoff. . Although we argue that this change in maternity leave takeup is relatively minor compared to the change in paternity leave, and at much higher margins, we might worry that it is partly the changed maternity leave that causes any changes in later labor market outcomes, not paternity leave.

To investigate this, we exploit the fact that some of the reforms expanded the paternity leave quota at the expense of maternity leave, while others lengthened the total leave. This means that we can exploit the stacked RD specification to get independent variation in the reform-induced shifts to both maternity and paternity leave use in a 2SLS setup:

$$\begin{aligned} y_{ir} &= \phi^{DL} L_i + \phi^{DM} M_i + \gamma_{r0} x_i \mathbb{1}(x_i < 0) + \gamma_{r1} \mathbb{1}(x_i \geq 0) + \epsilon_{ir} \\ L_{ir} &= \gamma_{LQ} Q_i + \gamma_{LS} S_i + g_{r0L}(x_i) \mathbb{1}(x_i < 0) + g_{r1L}(x_i) \mathbb{1}(x_i \geq 0)] + \eta^L_{ir} \\ M_{ir} &= \gamma_{MQ} Q_i + \gamma_{MS} S_i + g_{r0M}(x_i) \mathbb{1}(x_i < 0) + g_{r1M}(x_i) \mathbb{1}(x_i \geq 0)] + \eta^M_{ir} \end{aligned} \tag{5}$$

where L_i and M_i are paternity and maternity leave takeup for couple i. As before, we estimate this model separately for each parent and each relative time, stacking the reform-specific estimates as discussed in the main text. Notice that the variation in the two instruments are determined solely by the cutoff in birthdates, and that we have independent variation to separate the effects of both instruments because we stack all six reforms to parental leave. As before, we use local linear polynomials that are separate on either side of the cutoff and for each reform and a triangular kernel to control for the forcing variable. The outcome variable y_{ir} is labor market earnings.

When instrumenting for two endogenous variables in an IV-setup, it is not clear how to determine the optimal MSE-reducing bandwidth. We therefore use the same optimal bandwidths as for the separate RD estimates with a single endogenous regressor.

First stage results for the two endogenous variables are reported in Table A3. Notice that independent variation to identify both effects relies on stacking all reforms, so that we cannot perform these estimates separately by reform. Importantly, the reforms work exactly as we would expect: An increase in the daddy quota of 1 week increases paternity leave uptake by 0.9 weeks when we control for changes to the remaining quota for the mother, very similar to our baseline estimate of 0.86 weeks. Increasing the remaining leave for the mother (comprised of the maternal quota and the weeks of shared leave) increases maternity leave take up by 0.7 to 0.8 weeks. In contrast, the instruments do not work across spouses: Weeks of paternity leave quota does not affect maternity leave use when controlling for the remaining share available to the mother, in contrast to the balancing exercise in Table A2, while the remaining share for the mother does not affect leave uptake for the father when controlling for his own quota. Thus, the stacked specification where we instrument for both parents' leave take up circumvents the problem of the reforms affecting both margins of leave.

Table A3: First Stage Effects of Maternity and Paternity Leave Quotas

	Weeks of leave	
	Mother	Father
Paternity leave quota	0.014	0.90***
(Q_i)	(0.102)	(0.107)
Remaining leave for mother	0.768***	0.071
(S_i)	(0.151)	(0.159)
joint F	50.32	131.81
N	28,861	

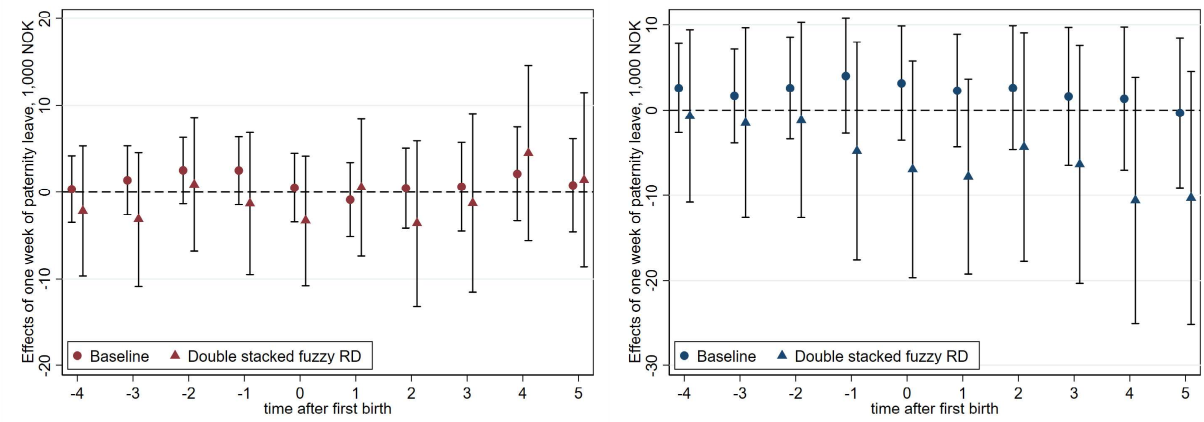
Notes: First stage results from stacked specification of all six parental leave reforms, instrumenting for weeks of paternity and maternity leave take up as described in equation 5. Bandwidths are the MSE-optimal bandwidths for the main specification with a single endogenous regressor. Reported first stage estimates are for the sample of mothers the year before birth. First stage estimates are very stable for fathers and other relative times, but may differ slightly due to unbalanced sample and different optimal bandwidths for mothers and fathers because of different variability in earnings. Standard errors clustered at the running variable, $'^{*}p < 0.1,^{**} p < 0.05.^{***}p < 0.01$ based on cluster robust, but not bias-corrected inference.

Results from the stacked fuzzy RD model where we instrument for both mothers' and fathers' leave take up is presented in Figure A4. The top panel presents effects of paternal leave on mothers' and fathers' earnings by child age, mirroring the estimates from the baseline model. For reference, the coefficients and confidence intervals from the stacked fuzzy RD model where we instrumented for paternity leave use only is added. The double IV model provides estimates that are well in line with the baseline model, confirming the precise zero effects of paternity leave on mothers' subsequent labor earnings. Just like in the basic model, it does not seem like paternity leave has a potential for reducing the child penalty.

Figure A4: Effects of Paternity Leave Use on Mother's and Father's Labor Earnings

(a) Paternity Leave on Mother's Earnings

(b) Paternity Leave on Father's Earnings



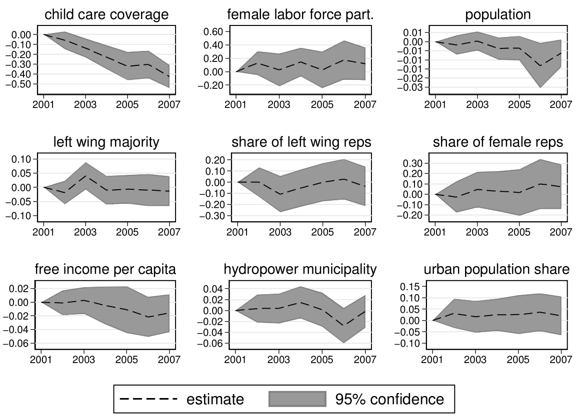
Notes: Figure shows the impact of a week of paternity leave use on mothers' (left) and fathers' (right) earnings over time, as estimated from a double IV stacked fuzzy RD as detailed in equation 5. For comparison we also show our stacked fuzzy RD estimates from the baseline model where we only instrument for weeks of paternity leave. 95% confidence intervals accounting for heteroskedasticity and clustering on the running variable indicated with whiskers around point estimates, bias-corrected confidence intervals in shaded bars.

The double IV specification inadvertently also estimates the effects of another week of maternity leave on parents' later earnings. Results are too imprecise to draw strong conclusions, but provide no evidence of any effects. In short, parental leave policies do not seem like a promising tool for reducing child penalties.

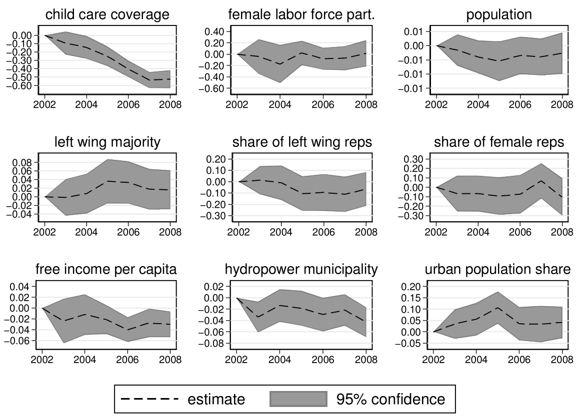
B Childcare: Robustness and Additional Results

Figure B1: Predicting Expansion of Slots from Pre-Reform Characteristics

(a) Childcare Coverage for 1-Year Olds

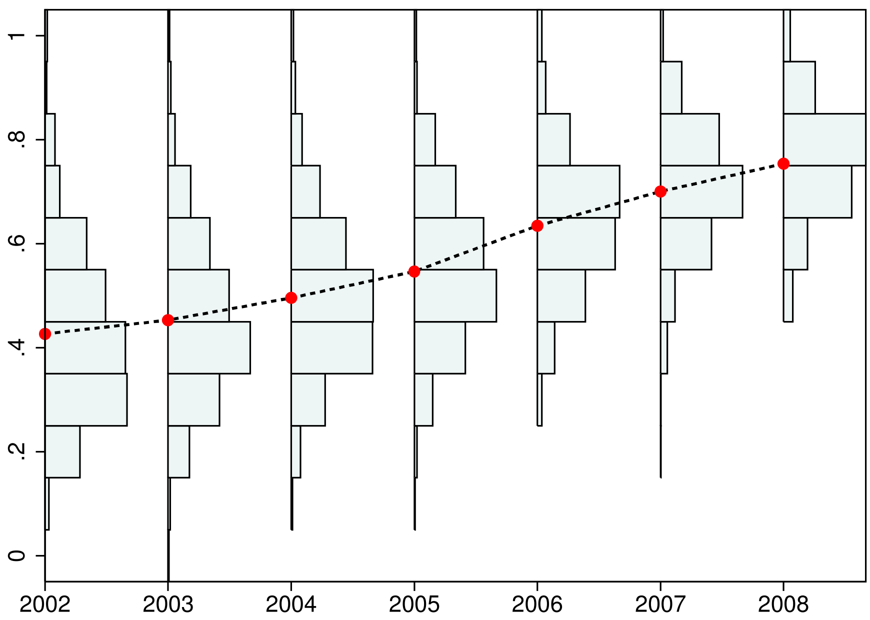


(b) Childcare Coverage for 2-Year Olds



Notes: Results from regression of our two instruments, childcare coverage at age 1 and 2, on municipality- and year fixed effects and an interaction of pre-reform characteristics interacted with year dummies, in a sample of municipalities over time. Plotted are the year-specific impact of the pre-reform characteristics on expansion of care in a particular year. 95% confidence intervals in grey, clustered at the municipality level.

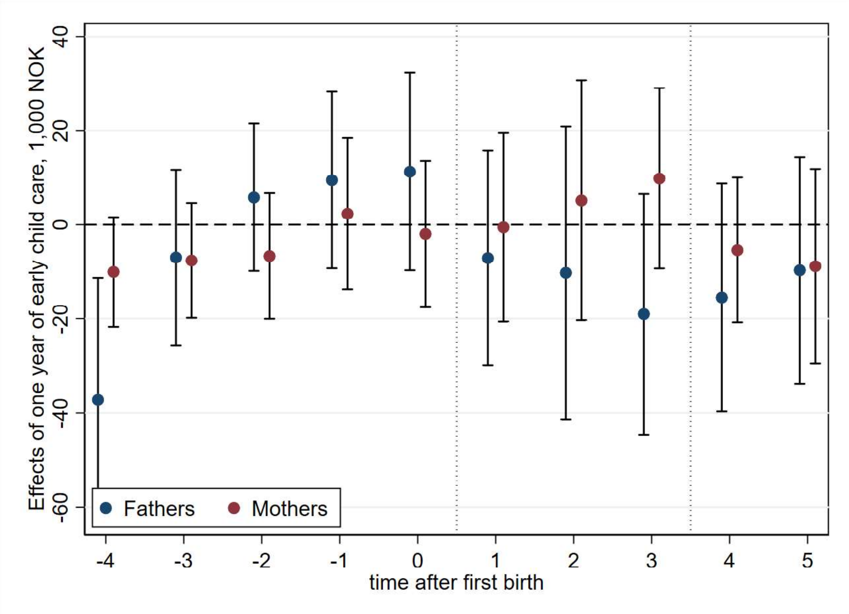
Figure B2: Distribution of Municipal Coverage Rates for 1-2 Year Olds



Notes: Figures created based on Statistics Norway Statistikkbanken, tables 09169 and 07459.

Figure B3 shows the impact of high quality, subsidized early childcare on each parent's earnings individually. Focusing first on the years of treatment, ages 1-3, we see that the estimates show point estimates of around 10,000-20,000 NOK for fathers at ages 1 to 3, where treatment happens, and NOK 5,000-10,000 for mothers. While these are not separately statistically significant, both of these contribute to decreasing the child penalty. The pre-birth outcomes, which we can think of as placebo outcomes, indicate small and insignificant impacts of future childcare use on past earnings for mothers, supporting the estimation strategy. Results for fathers are noisy, but suggest no impacts on earnings for the three years before birth.

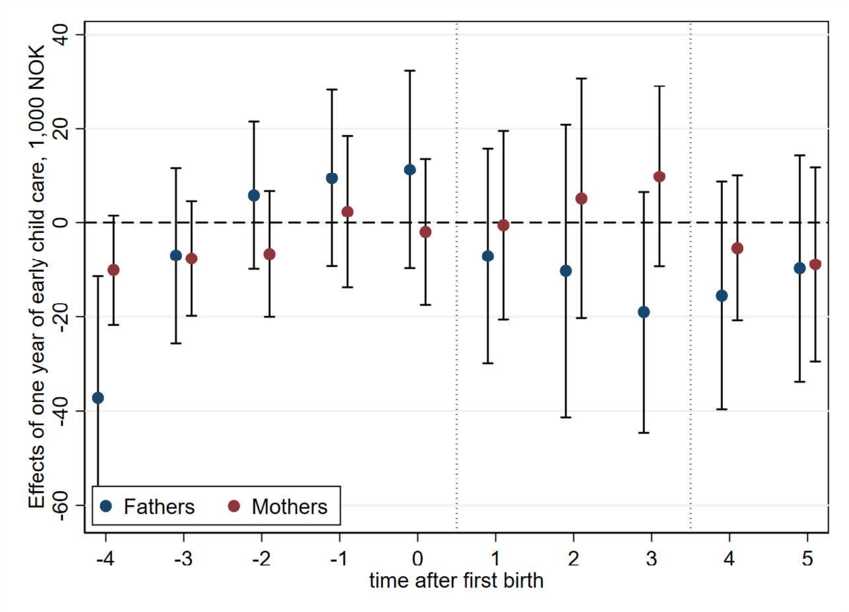
Figure B3: Impacts of Early Childcare Use on Mother's and Father's Earnings



Notes: Red circles indicate estimates of the impact of an additional year of early formal care use on mother's earnings over time. Blue circles show the same for fathers. 95% confidence interval, clustering at the municipality of residence at the time of treatment, indicated in whishers around the point estimates.

Figure B3 shows the impact of high quality, subsidized early childcare on each parent's earnings individually. Focusing first on the years of treatment, ages 1-3, we see that the estimates show point estimates of around 10,000-20,000 NOK for fathers at ages 1 to 3, where treatment happens, and NOK 5,000-10,000 for mothers. While these are not separately statistically significant, both of these contribute to decreasing the child penalty. The pre-birth outcomes, which we can think of as placebo outcomes, indicate small and insignificant impacts of future childcare use on past earnings for mothers, supporting the estimation strategy. Results for fathers are noisy, but suggest no impacts on earnings for the three years before birth.

Figure B3: Impacts of Early Childcare Use on Mother's and Father's Earnings



Notes: Red circles indicate estimates of the impact of an additional year of early formal care use on mother's earnings over time. Blue circles show the same for fathers. 95% confidence interval, clustering at the municipality of residence at the time of treatment, indicated in whishers around the point estimates.