Can the Child Penalty
Be Reduced? Evaluating Multiple Policy Interventions

Martin Eckhoff Andresen†    Emily Nix‡

Abstract

The arrival of children causes large earnings drops for mothers but not fathers, a stylized fact known as the “child penalty” that accounts for a substantial amount, and in some cases the vast majority, of remaining gender income gaps in wealthy countries. We evaluate two frequently suggested and implemented government policies aimed at reducing the child penalty: paternity leave and high quality child care. We find small and insignificant impacts of paternity leave use on both the child penalty and also on a new measure of father’s preferences for child care. In contrast, we find a 25% reduction in the child penalty from a large Norwegian reform that expanded access to publicly provided child care. These results suggest that while government interventions can reduce child penalties, providing viable alternatives to the mother’s time, such as high quality publicly provided child care, is more effective than paternity leave.

JEL-codes: I21, J13, J22, J71

Keywords: Gender wage gap, labor supply, child penalty, paternity leave, child care, event study, regression discontinuity, instrumental variables

†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, and Warwick University. We also thank Heather Antecol, Manuel Bagues, Sebastian Calonico, Matias D. Cattaneo, Nina Orangs, James Fenske, Yana Gallen, Trude Gunnnes, 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”. This version: November 22, 2021.

‡Corresponding Author: Emily Nix, Marshall School of Business, University of Southern California, enix@usc.edu; USC FBE Dept. HOH Hall – 231, MC-1422, 701 Exposition Boulevard, Ste. 231 Los Angeles, CA 90089-1422 USA
1 Introduction

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

Figure 1 shows our estimates of the impacts of children on earnings of mothers and fathers over time in Norway. The figure suggests the child penalty in Norway has decreased significantly over time. Mothers used to experience a 50% drop in average earnings following the arrival of their first child in the 1970’s, which has decreased to less than a 30% drop in earnings more recently. In contrast, fathers used to experience what appears to be a “child premium”, approximately a 10% increase in earnings in the early 1970s, which has reduced to no impact of children on earnings more recently. During this period the Norwegian government adopted a number of programs that may have played a role in this reduction.

In this paper we examine two prominent policies and their impacts on the child penalty: paternity leave and government subsidized child care. As a means of increasing fathers’ involvement in raising children, the so called daddy quotas (leave that can only be taken by fathers) of the Scandinavian countries have attracted considerable interest. Similarly, there has been significant discussion of the impact of subsidized child care on labor market outcomes of new mothers.

---

1 Lundberg and Rose (2002) find that men’s labor supply and wage rates used to actually increase following birth of children, which is consistent with the historical results in Figure 1 although this “fatherhood premium” has dissipated over time. See also Choi et al., 2008.

2 Kuziemko et al. (2018) suggests that women may not anticipate this large cost of children in terms of their labor earnings. This may have important implications for women’s demand before giving birth for policies that could reduce the child penalty, and suggests that studying demand for such programs before and after the arrival of children could be a productive avenue for future research, to better understand how preferences evolve over time. See also Wiswall and Zafar, 2021.

3 See Carneiro et al. (2015), Dahl et al. (2014), and Havnes and Mogstad (2011) for studies of the impact of early introductions of maternity leave, paternity leave, and child care, respectively.

4 Another prominent policy is paid maternity leave, but the introduction of maternity leave in Norway occurred in 1977. This is too early for us to examine its effect on the child penalty using our data. However, to some extent the extensive margin impacts of maternity leave are less interesting, as almost all countries have already chosen to implement it, with the notable exception of the United States.

5 In addition to Scandinavian countries, a number of other countries have introduced similar quotas, including Ireland (14 weeks), Slovenia and Iceland (13 weeks), Germany (8 weeks), Finland (7 weeks), and Portugal (6 weeks) (OECD, 2014).
There are a number of theoretical reasons why these two policies could reduce the child penalty. Paternity leave could increase the utility fathers get from spending time with children and change gender norms around child care, with possibly long term effects after the period of leave. Subsidized early child care reduces the cost of market care in the household budget constraint and could increase labor supply of the mother if she is the counterfactual caretaker. Whether these policies work in practice is an empirical question. We use administrative data from Norway to estimate the causal impact of each policy in the same setting, allowing us to compare and contrast how each impacts the child penalty. Our results are particularly important given that a number of other countries are currently considering adopting or expanding similar policies.

To identify the causal impact of paternity leave on the child penalty we use a stacked regression
discontinuity design to estimate the impact of six reforms to the paid paternity leave quota in Norway from 2005 to 2014. We estimate a strong first stage: the reforms significantly increase paternity leave takeup. Despite showing that fathers do take the additional leave, 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. While we report null findings on labor supply, paternity leave could instead impact norms within the couple in ways not picked up by labor market earnings. We propose fathers’ total takeup of parental leave for subsequent children as a measure of changes in the norms around child care within the couple. We find that paternity leave does not impact the distribution of shared leave for future children. Combined, these results indicate that paternity leave has limited potential to reduce the child penalty and does not substantially change norms within couples, at least for the dimensions we examine.

To identify the causal impact of access to high quality child care on the child penalty we use a large-scale Norwegian reform from 2002 that expanded child care availability for 1- and 2-year-olds. The reform increased subsidies to child care 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 child care 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 positive effects on mothers’ labor income at ages two and three that scales to reduce the child penalty by around 25 percent for each year of early child care use. Although impacts are not persistent in the long run, the sizable short run effects suggest that subsidized child care 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. [Andresen and Nix 2019] shows that there is a large child penalty in our setting, Norway, with new mothers experiencing a reduction in labor market earnings of 23 percent compared to no impact on new fathers. These results are consistent with many papers documenting large effects in other countries, such as 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. In this paper we provide evidence on how successful government policy might be in reducing 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 economics consequences of family leave policies. We contribute to the literature by isolating the impact of these policies on the child penalty and estimating the impacts of both policies in a single setting. Our results on paternity leave are related to and consistent with [Antecol et al. (2018)] who find that moving toward more gender neutral benefits in response to children does not help women in academia, and may even hurt their careers relative to men. We show that the results from [Antecol et al. (2018)] are not unique to academia. These results also tie in to a larger literature that examines the impacts of paternity leave on parents’ earnings and labor supply, finding mixed results. Furthermore, we show no effect of exposure to paternity leave for the first child on leave use for subsequent kids, suggesting that preferences for leave taking are not substantially affected by exposure to paternity leave. This result is similar to the finding in [Bana et al. (2018)], that men take much less paid family leave than women in California. While we find no impact on the child penalty or future leave taking of fathers, this does not rule out other positive impacts of paternity leave. Regarding child care, 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 child care availability. Of most relevance here are [Havnes and Mogstad (2011)] who find small effects from a child care reform for preschoolers, and [Andresen and Havnes (2019)] who find considerably larger effects from a child care reform for toddlers. In this paper we focus on the impact of these policies on one particular outcome of interest, the child penalty.

---

For a good overview of this literature, see [Rossin-Slater (2017)]. Most closely related to this paper, [Rege and Solli (2013)] find a decrease in fathers’ earnings long term in Norway from a 1993 reform using a difference in difference approach. [Druedahl et al. (2019)] find that a Danish increase in the daddy quota from 2 to 4 weeks increased mothers’ share of household earnings; [Johansson (2010)] finds that a Swedish policy increased mother’s earnings but had no impact on fathers; [Ekberg et al. (2013)] find that fathers are no more likely to take sick leave to care for a sick child long term using a Swedish reform; [Cools et al. (2015)] estimate the effect of paternity leave extensions in Norway and, like us, find no effect on traditional labor supply allocations in the family although they do find improvements in children’s test scores; and [Andersen (2018)] find 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 and [Persson and Rossin-Slater (2019)] find that when fathers have more flexibility to stay home, there are positive impacts on the mother’s health.
Parental Leave and Child Care in Norway

Following a birth, Norwegian parents have been entitled to generous paid parental leave since 1977. Total parental leave is currently 49 weeks at 100 percent replacement or 59 weeks at 80 percent replacement rate. The length of leave has steadily increased since the mid 1980s. In Figure 2a we report every leave reform in Norway from 1992 to 2014. The maternal and paternal quota columns report the amount of parental leave in weeks that is reserved exclusively for the mother and father. The remaining leave can be shared among parents however they choose and is reported in column 5. Leave benefits are capped at around 640,000 Norwegian kroner (NOK, 2021), roughly 75,000 USD, with many employers topping up.

In order to qualify for leave, a parent must have been employed for at least 6 of the 10 months prior to birth, and the 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 are entitled to a one-time-benefit of 90,300 NOK or approximately 10,000 USD. In addition to paid leave, all parents have job protection for another year if they want to take additional unpaid leave.

Following parental leave, Norway has a well-developed, well-regulated, and highly subsidized child care sector. Because of the heavy subsidies for formal care, the market for paid child care outside this system is very small, but subsidies are available for both private and public suppliers of formal care. For children not in formal care and aged 13 to 35/36 months, depending on the cohort, there was a monthly cash-for-care benefit.

Figures 3b-3c show the child care coverage rates over time in Norway, separately by age of the children. These figures show that the formal care sector for preschoolers was well developed in Norway by the early 2000’s, with more than 80 percent of Norwegian 4 and 5 year olds attending care. For toddlers (below 3 years), however, coverage was much lower (between 30 to 50 percent), and the market was strongly rationed. These facts are documented in greater detail in Andresen and Havnes (2019), including additional evidence from surveys on the actual and preferred modes of child care for children at these ages.

---

7 Throughout 2001-2009 the benefit was around 3,500 NOK nominally, equivalent to approximately or 550 USD in 2021.
8 The prevalence of care is the result of a reform and gradual expansion of formal care for these children in the 1970’s Havnes and Mogstad (2013).
The underrepresentation of children below 3 in formal care was the impetus for the Child Care Concord in 2002, a broad, bipartisan agreement to increase the availability of care for toddlers. Following this reform, coverage increased rapidly for 1- and 2-year-olds over the next years as shown in Figure 3b. This expansion varied considerably between municipalities and over time as illustrated in Figure 3c, making the expansion of care availability a potential instrument for the endogenous choice of how much child care to use. This is the variation we use for identification in this paper.

Figure 2: Parental Leave and Child Care Coverage Reforms

(a) Parental leave reforms in Norway, in weeks

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

(b) Child care coverage rates

Subfigure (a) shows parental leave in weeks. Weeks of leave at 80% replacement rate in parentheses. See NOU 2017:6 (2017).

For Subfigures (b) and (c) the childcare coverage is taken from Statistics Norway Statistikkbanken, tables 09169 and 07459.
3 Data, Samples and Summary Statistics

We use data from Norwegian administrative registers covering the entire resident population, for the years 1967-2018. Our main outcome of interest for both applications is the child penalty, estimated from annual labor market earnings (including taxable benefits) of fathers and mothers. To measure parental leave use, we use data from the Norwegian Public Insurance system (FD-Trygd) for information on all leave spells. The data does not contain direct links to the child for whom the leave is taken, only to the individual who takes leave. We therefore infer the relevant child from the birthdates of the children. 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 the rules for parental leave, which can be taken up to the age of three, but any remaining leave not taken by the time the next child is born 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 measure time taken away from work on paid leave.

For all applications, we restrict the samples to first-time parents. For the child care analysis, we start with all children born in the years 2000-2006, who will be subject to the reform-induced expansions of care in 2002-2008. We assign children to their municipality of residence at the age of 1 and look at couples where both parents reside in that municipality when the child is 1. While much of the literature restricts the sample to children without younger siblings, we view future fertility as a potentially endogenous outcome of the reform, and therefore do not restrict the sample to youngest children. This leaves us with a sample of around 103,000 couples.

For the parental leave analysis, we restrict the sample to first time parents in the reform years 2005, 2006, 2009, 2011, 2013 and 2014. We include in the sample only couples where the mother took

---

9 There are five types of leave spells recorded: regular parental leave spells; pregnancy leave spells available for mothers with jobs that impose health risks to the unborn child; adoption leave spells; combined leave spells; and other leave spells. More than 97 percent of the leave spells recorded are for regular leave spells, and we focus on these.

10 In cases of graded leave (working part-time) we compute the efficient days at home for each leave spell. 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 in our sample) to 60 weeks.

11 This includes a few thousand twins. Clustering at the municipality level accounts for within-family clustering.

12 Notice that because we restrict the sample to first born children, it is a little less than half the size of the samples of cohabiting mothers and fathers in [Andresen and Havnes (2019)]. Because of our focus on labor supply over time, we also measure child care use throughout the full 13 - 36 months period we can measure.
some leave, which means that the father is also eligible.[13] As described in Section 4, these samples are further refined by using optimal bandwidths around the timing of the reform. This leaves is with a potential sample of 93,123 couples. For both applications, we construct a panel of labor market earnings from 4 years before birth to 5 years after birth. This means that the paternity leave sample is slightly unbalanced for the final reform in 2014 because we cannot measure earnings longer than 2018.

4 Identifying Child Penalties and Policy Impacts

4.1 Identifying Child Penalties Using an Event Study

In order to identify the child penalty, we estimate a standard event study model

\[
y_{it} = \sum_{j \neq -1} \beta_j [t - E_i = j] + \sum_{j \neq -1} \alpha_j [t - E_i = j, m(i) = 1] + \theta_m(i) a(it) + \gamma_m(i) t + \eta_m(i) + \epsilon_{it},
\]

where \( m(i) \) is a dummy equal to 1 for mothers and 0 for fathers. \( 1[A] \) is the indicator function for event \( A \). \( \alpha_j \) captures the child penalty or the difference in the impact of children on mothers versus fathers. In some cases (Figure 1 and some Appendix figures), we report the impact of children on mothers and fathers separately. For these exercises, \( \beta_j \) captures the earnings dynamics around birth for fathers and \( \alpha_j + \beta_j \) captures earnings dynamics for mothers. As discussed in [Andresen and Nix (2019)], the assumptions required for \( \beta_j \) and \( \alpha_j + \beta_j \) to be interpreted causally are less likely to be met given the pervasive presence of pre-trends in the event studies looking at the impact of children on mothers’ and fathers’ earnings separately. In contrast, in order to interpret \( \alpha_j \) 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. Pre-trends support this assumption.

We also control flexibly for yearly shocks (\( \gamma_m(t) \)) that are allowed to vary freely for both mothers and fathers and flexible age profiles \( \theta_{ma} \) that account for the fact that most couples tend to have their first child at times when earnings dynamics are increasing. Standard errors are clustered by couple

[13] footnote describing why this is likely not endogenous.
and robust to heteroskedasticity.

We scale the absolute estimates to percentage effects following Kleven et al. (2019b) to construct the following scaling factors:

\[ S_{jm} = \mathbb{E}(\hat{\theta}_{a(it)m} + \hat{\eta}_m + \gamma_{ma(it)} | E_i - t = j, m(i) = m). \] (2)

Thus, we scale the child penalty with the mother’s expected earnings, so \( C_j = \frac{\alpha_j}{S_{jm}} \). In practice, the denominator is simply a weighted average of the yearly shocks and age profiles for mothers. \(^{14}\) \( C_j \) is thus the child penalty in at event time \( j \) relative to the predicted outcome of the mother absent children. We use a bootstrap to calculate standard errors for the scaled estimates, clustering at the couple.

When estimating the impact of the two reforms, we will focus on the impacts on the scaled child penalty.

### 4.2 Identifying the Effects of Paternity Leave Using Regression Discontinuities

We use a regression discontinuity design to identify the impact of paternity leave on the child penalty. Specifically,

we use the 2005, 2006, 2009, 2011, 2013, and 2014 paternity leave reforms in our analysis. \(^{15}\) Our identification strategy exploits the fact that parents of children born just before the reforms were not subject to the changes in parental leave quotas, whereas parents of children born right after each reform were subject to the changes. The reforms were generally announced in October the year before implementation as part of the budgeting process, making it nearly impossible to plan conception in response to the announcement of the quota change in order to manipulate birth dates around the cutoff in July. Appendix Figure A2 verifies that there is no statistically significant change in the density of births around the cutoff for any of the reforms.

In order to increase precision, in our main results we combine the six reforms, although we also

---

\(^{14}\)This allows the scaling factor to increase over time as age profiles grow.

\(^{15}\)We exclude the 1992 reform because it requires a donut-RD framework to identify the effects, which is not necessary for the other results.
present estimates by each year in the appendix.\[16\] The common way of stacking multiple reforms in RD studies is to re-center the running variable to be zero at the relevant cutoff for all individuals and run semi-parametric 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. In addition to this restriction, naive re-centering is problematic in our case because the treatment scaling varies across reforms, from a decrease of four weeks to an increase of four weeks and various changes in between. An alternative and more straightforward way to stack the estimates is to allow the local polynomials of the running variable to vary by cutoff and use the cutoff-specific optimal bandwidths and kernel weights from the individual specifications. The results are scaled to reflect one week of quota expansion by using an indicator of the number of weeks of quota increase rather than a dummy at the cutoff. Unfortunately, we cannot calculate bias-corrected standard errors for this specification, but we argue that the problem should be relatively minor.\[17\]

Formally, we estimate:

\[
y_{it} = \beta_t I_i + f_{tc(i)}(x_i) + \mu_{it}
\]
\[
L_i = \sum_c \gamma_g [ (x_i \geq 0) Q_c + g_c(x_i) ] + \xi_{it},
\]

where \(x_i\), the running variable, is the number of days after the reform date that the child was born. \(f_{tc(i)}(x_i)\) and \(g_c(x_i)\) are local linear polynomials that are separate on either side of the cutoff. 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 in order to obtain estimates local to the cutoff. We estimate and report robust bias-corrected confidence intervals (Calonico et al. 2014) together with the conventional, heteroskedasticity-robust confidence intervals. We then scale the effects on earnings to reflect the percentage changes in the child penalty.\[18\]

---

\[16\] Appendix Figure A4 reports estimates using the reforms separately, which are consistent with our main findings.\[17\] First, notice that the difference between the conventional and the robust standard error estimate for the reform-specific cutoffs in Table 1 are small, indicating that the variance coming from the approximation error is relatively minor. Second, the approximation error should be smaller for the stacked than the alternative naive pooled estimator because we allow the local polynomials to differ between cutoffs and thus approximate the unknown functions better. Nonetheless, inference from this specification is only correct if the model is well specified, so that approximation error vanishes asymptotically.\[18\] Many models in this section are estimated using the robust RD commands for Stata written by Matias D. Cattaneo et al., to whom we owe thanks. These include rdrobust, rddensity and rdbwselect. See Calonico et al. (2018) and Cattaneo et al. (2018).
For the stacked fuzzy RD, we revert to the cutoff-specific treatment indicators as instruments because the fuzzy RD takes care of scaling. This specification reproduces the cutoff-specific first stage estimates for each reform reported in Appendix Figure A1 for a given bandwidth and so is a natural way to stack the reforms. When interpreting these fuzzy RD estimates, it is important to keep in mind that these estimates are local average treatment effects: they capture the effects of additional leave use on earnings for couples induced to use more leave because they were exposed to the reforms. In our case, the compliers represent *unwilling users* of paternity leave, because these couples were free to distribute more leave than the quota to the father irrespective of the reform (see column 5 of Figure 2a). In case of heterogeneous treatment effects, the average effect for the compliers need not be the same as that for the population. Despite this, we argue that the LATE is a particularly policy-relevant treatment effect in our case because it reflects the effects of paternity leave use for fathers induced to take more leave by the policy instrument, which is arguably the population of interest to policy makers.

The critical assumption for the validity of our RD approach is that the underlying regression functions are continuous at the threshold. This implies that the population of couples around the discontinuity are identical. We provide empirical support for this assumption using balancing tests in Appendix Table A2.

### 4.3 Identifying the Effects of Early Child Care Using Instrumental Variables

To identify the impact of early child care on the child penalty we take our baseline event study specification and see how adding the measure of individual early child care use affects the child penalty. Because child care is endogenous to labor supply, we instrument care use with the expansion of slots for 1-year olds at age 1 and for 2-year olds at age 2 in the following IV model:

---

19 An important imbalance revealed in this table is maternity leave take-up, as some of the reforms we exploit increase paternity leave quota at the expense of the shared leave most often taken by the mother. We do not believe these relatively small changes in maternity leave take-up from already high levels to be driving our results. In Appendix A.4 we exploit the fact that some of these reforms expanded paternity leave use at the expense of maternity leave, while others expanded the total leave length. This allows us to instrument for both the maternity and paternity leave use, confirming the baseline results of the effects of paternity leave on the child penalty.
\[ y_{it} = \pi_k + \gamma_{T_{it}} + \beta_{a_{it}} + \phi_t m_i + \epsilon_{it} \]

\[ m_i = \tilde{\pi}_k + \tilde{\gamma}_{T_{it}} + \tilde{\beta}_{a_{it}} + \gamma_1 CC^1_k + \gamma_2 CC^2_k + \tilde{\epsilon}_{it} \]  \hspace{1cm} (4)

where \( \gamma_{T_{it}} \) are calendar year fixed effects, \( \pi_k \) are municipality fixed effects, \( \beta_{a_{it}} \) are age fixed effects for the parent (in years) and \( m_i \) is our measure of child care use from ages 13 to 36 months from the cash for care data. The instruments are \( CC^1_k \), the share of slots for 1-year-olds in the municipality at age 1 to the population of 1-year-olds, and \( CC^2_k \), the same share for 2-year-olds, measured at the relevant age of the child.

The variation we exploit thus comes from the 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 parents’ child penalty, our approach recovers the causal effect of an extra year of early child care on the child penalty for the compliers: the mothers who take up the newly expanded slots. Because child care was strongly rationed before the reform, it is natural to think of the compliers as the mothers of children who wanted child care before the reform, but were restricted by the low supply. [Andresen and Havnes (2019)] shows that 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 some support for the idea that expansions did not systematically vary across municipalities with different pre-reform characteristics (except, of course, the initial coverage rate), while [Andresen and Havnes (2019)] provide a range of specification checks that demonstrate the robustness of the instrument.

5 Child Penalties Over Time and the Impacts of Paternity Leave and Early Child Care

As we can see from Figure 1, there has been a reduction in the child penalty over time, which is consistent with the estimated impacts of children on individual earnings dynamics shown in Figure 1. With a better understanding of the impact of children on families over time, we now turn to the
impacts of paternity leave and expanded access to child care on these child penalties.

We start by reporting first stage estimates for both policies in Table 1. In Panel A we see clear effects of the paternity leave reforms on the take up of paternity leave. The first stage estimates are always significant, whether using robust bias-correcting inference or conventional inference that only accounts for heteroskedasticity. First stage estimates from equation 4 are presented in Table 1 Panel B, column 1, and show that expansions of care both at age 1 and at age 2 have a strong impact on early child care use, with an additional slot in care at age 1 increasing care use by around 0.8 years and at age 2 by about 0.6 years. Because our endogenous variable captures the intensity of use throughout the full period, these coefficients are not 1; as additional slots are generally opened in August, children may not have the chance to exploit them to capacity the whole year. The IV strategy thus scales the reduced form estimates to reflect a full year of early child care use. The $F$-statistic is above 150, indicating a very strong first stage.

---

20 Appendix Figure 3
21 Note that all results have been scaled to reflect one week of quota expansion.
Table 1: First Stages for Parental Leave and Child Care Coverage Reforms

<table>
<thead>
<tr>
<th>Panel A: First stage estimates of paternity leave</th>
<th>Pooled</th>
<th>Stacked</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD estimate per week</td>
<td>0.86***</td>
<td>0.88***</td>
</tr>
<tr>
<td>conv. standard error</td>
<td>(0.10)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>robust standard error</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>conventional p-value</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>robust p-value</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>93,123</td>
<td>93,123</td>
</tr>
<tr>
<td>Optimal bandwidth</td>
<td>60.6†</td>
<td>60.6†</td>
</tr>
<tr>
<td>Efficient observations</td>
<td>31,584</td>
<td>31,584</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: First stage estimates of child care reform</th>
<th>Years of child care use age 1-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage rate at age 1</td>
<td>0.787***</td>
</tr>
<tr>
<td></td>
<td>(0.0543)</td>
</tr>
<tr>
<td>Coverage rate at age 2</td>
<td>0.630***</td>
</tr>
<tr>
<td></td>
<td>(0.0629)</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>✓</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>✓</td>
</tr>
<tr>
<td>Age profiles</td>
<td>✓</td>
</tr>
<tr>
<td>Education-specific age profiles</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>103,172</td>
</tr>
<tr>
<td>mean dep. var</td>
<td>1.031</td>
</tr>
<tr>
<td>$F$</td>
<td>167.9</td>
</tr>
</tbody>
</table>

Source: Panel A shows robust 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 refer to the weighted average of reform-specific estimates. Stacked estimates stacks models for each reform, allowing polynomials to vary over cutoffs and using the cutoff-specific bandwidths and weights. Conventional standard errors are heteroskedasticity-robust, but not bias-corrected. ***p < 0.01, **p < 0.05, *p < 0.1, using conventional, heteroskedasticity-robust standard errors. †indicates average bandwidth. Panel B shows the first stage estimates for the childcare reform, from equation (4). Standard errors in parentheses, clustered at municipality.

Figure 3 reports the impacts of paternity leave on the child penalty using the stacked and pooled fuzzy RD estimates. The $y$-axis in this figure represents the percentage change in the child penalty. Point estimates are close to zero, suggesting no impact of paternity leave on the child penalty. This zero is relatively precise, as the lower bound of the confidence intervals rules out reductions larger than 22%

---

22 Effects on mothers' and fathers' annual incomes are reported in Appendix Figure A3.
around 5 to 7 percent of the child penalty per week of paternity leave use for children ages 1 through 5.\footnote{One might believe that paternity leave could have long run effects on norms. However, the first paternity leave reform occurred in 1992 in Norway, which begs the question, how long should one wait to see long run effects?}

While our main results suggest paternity leave does not substantially reduce the child penalty, such leave might influence gender norms or preferences around the distribution of home work without affecting labor market earnings. One possible measure of such norms is increased use of shared leave by fathers for future children. To investigate whether paternity leave impacts the father’s choice to spend time with his children, we exploit the fact that many of the fathers that have a child around the time of the reforms subsequently go on to have more children. We estimate our fuzzy RD model using as an outcome the father’s 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).\footnote{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, but Farré and González (2019) find negative impacts of paternity leave on fertility in Spain.}

In order not to use the outcome variable for one child as the treatment variable for another, we restrict attention to the first child each father has in one of the reform years and look at outcomes for the next child. Notice that we cannot use the 2014 reform for this exercise, as we cannot reliably measure paternity leave use for kids born after 2014.

Figure 4c provides the results of this exercise for the pooled and stacked estimates for all reforms. There is little evidence of any permanent impact on norms as measured by take-up of paternity leave for later kids. Moreover, none of the individual reforms provide statistically significant results and point estimates are negative except for the 2013 reform, where the efficient sample size is only 159 children and we find a marginally significant effect, (see Appendix Table A3). Focusing on our preferred stacked estimates in the figure, the results indicate a non-significant effect of .1 less weeks of leave for subsequent children for each week of paternal leave taken for the first child, where the top of the 95 percent confidence interval rules out effects larger than around 0.12 week extra leave for subsequent kids per week of leave for the first child.

Figure 3b reports estimates of the impacts of a full year of child care use on the child penalty. Results show that the child penalty is reduced by around 25 to 30 percent for mothers when their children are between the ages 2 and 3, but the impact appears only in the years of treatment (note that age 3 is included in treatment since some children will receive slots just before turning 3, so treatment will also...
occur at age 3 for these kids). Appendix Figure B2 presents results separately for mothers’ and fathers’ earnings. These results show that the main impact is on mothers, who see significant increases in their labor market earnings. As a robustness check, we include the education level-specific age profiles in equation 4. The first stage from this specification is virtually identical (see Table 1, Panel B, Column 2). The second stage results are also very similar.

We conclude that early child care shows more promise as a policy tool for reducing child penalties than paternity leave, although it does not appear to have a permanent impact. However, the impact of paternity leave is measured per week of leave added while the impact of child care is for a full year of child care. While these are usually the relevant marginal increases for each program, policy makers may wish to know the impact per dollar spent. To that end, we find that paternity leave point estimates suggest on average a statistically insignificant 1.1 percent reduction of the child penalty per 10,000 NOK spent (about 1,100 USD) and based on the confidence interval we can rule out impacts larger than 6.7 percent per 10,000 NOK. In contrast, child care point estimates suggest on average a statistically significant 1.7 percent reduction of the child penalty 10,000 NOK, and based on the confidence intervals the impact could be as high as 3.7 percent per 10,000 NOK spent.

---

25 For these back of the envelope calculations, public cost of child care is estimated at around 160,000 NOK per year in subsidies (Andresen and Havnes 2019), while the cost of paternity leave is estimated as the average wage of the fathers in our sample, capped at 6G, to represent the replacement rate of earnings.
Figure 3: Impacts of paternity leave and formal childcare on the child penalty and norms

(a) Fuzzy RD estimates of paternity leave use child penalty

(b) Impact of a year of early child care use on the child penalty

Reform year Pooled Stacked

RD estimate per week -0.268 -0.092
conv. standard error (0.359) (0.11)
robust standard error 0.412

conventional p-value 0.428 0.40
robust p-value 0.491

Observations 53,456 53,456
Efficient observations 16,896 16,896

(c) Paternity norms: Fuzzy RD of paternity leave on leave for next child

Notes: Panel (a) shows fuzzy RD estimates of the impact of an additional week of paternity leave use on the mother’s child penalty, using all six reforms. The pooled estimate refers to the weighted average of the reform-specific estimates, while the stacked estimate stacks the cutoff-specific specifications for precision. Robust bias-correcting inference reported for the pooled estimate and conventional, heteroskedasticity-robust inference for the stacked estimate. Panel b) shows effects of early use of paternity leave on use of leave for later children. Panel (c) shows IV results from equation (4) reflecting the impact on mother’s labor earnings in 1,000 NOK across child age for an extra year of early child care use at ages 13-36 months, scaled with the estimated child penalties from the first part of the paper to represent the change in the child penalty for mothers.

6 Conclusion

In this paper we examined two policies that might reduce the child penalty: paternity leave and subsidized early child care. 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. In addition, paternity leave has no impact on whether the father takes additional leave for future children, pointing to limited impact on gender norms.
While there may exist other benefits of paternity leave, for example, Persson and Rossin-Slater (2019) find positive impacts on maternal health, if the goal of paternity leave is to reduce the gender income gap, the results in this paper suggest that this is a very expensive program that does not achieve its objective.

In contrast, we show that early child care use reduces the child penalty by around 25 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 child care to families, not on offering paternity leave to fathers. While we have focused on two of the most commonly proposed policies to reduce the child penalty, there are a number of additional policy changes that could impact the child care penalty differently. Better understanding the impacts of these other possible policies is a productive avenue for future research.

References


Online Appendix

A Paternity leave: Robustness and Additional Results

A.1 First Stage Graphically and Table with All Reforms Individually

Figure A1: Fuzzy RD first stage estimate

Note: 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.
Table A1: RD first stage estimates separately by year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RD estimate per week</td>
<td>0.79**</td>
<td>1.05***</td>
<td>0.98***</td>
<td>0.82***</td>
<td>0.69***</td>
<td>0.81***</td>
<td>0.86***</td>
<td>0.88***</td>
</tr>
<tr>
<td>conv. standard error</td>
<td>(0.33)</td>
<td>(0.33)</td>
<td>(0.095)</td>
<td>(0.28)</td>
<td>(0.23)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>robust standard error</td>
<td>0.40</td>
<td>0.40</td>
<td>0.11</td>
<td>0.33</td>
<td>0.28</td>
<td>0.14</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>conventional p-value</td>
<td>0.016</td>
<td>0.002</td>
<td>0.000</td>
<td>0.003</td>
<td>0.003</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>robust p-value</td>
<td>0.031</td>
<td>0.007</td>
<td>0.000</td>
<td>0.006</td>
<td>0.019</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>14,598</td>
<td>15,111</td>
<td>16,501</td>
<td>16,500</td>
<td>16,173</td>
<td>14,240</td>
<td>93,123</td>
<td>93,123</td>
</tr>
<tr>
<td>Optimal bandwidth</td>
<td>64.7</td>
<td>54.1</td>
<td>73.2</td>
<td>47.8</td>
<td>65.3</td>
<td>42.5</td>
<td>60.6†</td>
<td>60.6†</td>
</tr>
<tr>
<td>Efficient observations</td>
<td>5,302</td>
<td>4,797</td>
<td>6,877</td>
<td>4,659</td>
<td>6,264</td>
<td>3,685</td>
<td>31,584</td>
<td>31,584</td>
</tr>
</tbody>
</table>

Notes: Robust 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 refer to the weighted average of reform-specific estimates. Stacked estimates stacks models for each reform, allowing polynomials to vary over cutoffs and using the cutoff-specific bandwidths and weights. Conventional standard errors are heteroskedasticity-robust, but not bias-corrected. **p < 0.01, *p < 0.05, †p < 0.1, using conventional, heteroskedasticity-robust standard errors. †Average bandwidth.

A.2 Balancing Tests

Table A2 provides sharp RD balancing tests for a range of covariates in the baseline RD model. Figure A2 provides robust local polynomial estimates of the density of births around the cutoff. Reduced form and first stage estimates separately by reform is plotted in Figure A4.
Table A2: Sharp RD balancing tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother’s age</td>
<td>RD estimate</td>
<td>0.030</td>
<td>-0.19</td>
<td>0.030</td>
<td>-0.089</td>
<td>0.088</td>
<td>-0.032</td>
<td>-0.027</td>
<td>-0.0043</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.23)</td>
<td>(0.25)</td>
<td>(0.066)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.072)</td>
<td>(0.067)</td>
<td>(0.042)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.80</td>
<td>0.44</td>
<td>0.77</td>
<td>0.41</td>
<td>0.43</td>
<td>0.80</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Father’s age</td>
<td>RD estimate</td>
<td>0.18</td>
<td>0.17</td>
<td>0.020</td>
<td>-0.20</td>
<td>0.19</td>
<td>0.0072</td>
<td>0.075</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.25)</td>
<td>(0.26)</td>
<td>(0.069)</td>
<td>(0.16)</td>
<td>(0.16)</td>
<td>(0.080)</td>
<td>(0.079)</td>
<td>(0.046)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.51</td>
<td>0.50</td>
<td>0.97</td>
<td>0.31</td>
<td>0.26</td>
<td>0.82</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Maternity leave</td>
<td>RD estimate</td>
<td>0.99*</td>
<td>0.10</td>
<td>-0.33***</td>
<td>-0.14</td>
<td>-0.059</td>
<td>-0.67****</td>
<td>-0.028</td>
<td>-0.36***</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.49)</td>
<td>(0.49)</td>
<td>(0.11)</td>
<td>(0.23)</td>
<td>(0.20)</td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.069)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.046</td>
<td>0.96</td>
<td>0.013</td>
<td>0.64</td>
<td>0.98</td>
<td>0.000</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Father’s ed. years of ed.</td>
<td>RD estimate</td>
<td>-0.0062</td>
<td>-0.32</td>
<td>0.0092</td>
<td>-0.050</td>
<td>0.017</td>
<td>-0.012</td>
<td>-0.055</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.20)</td>
<td>(0.20)</td>
<td>(0.047)</td>
<td>(0.074)</td>
<td>(0.091)</td>
<td>(0.054)</td>
<td>(0.046)</td>
<td>(0.029)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.75</td>
<td>0.12</td>
<td>0.92</td>
<td>0.60</td>
<td>0.84</td>
<td>0.94</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Mother’s ed. years of ed.</td>
<td>RD estimate</td>
<td>0.055</td>
<td>-0.26</td>
<td>0.050</td>
<td>-0.080</td>
<td>0.098</td>
<td>0.040</td>
<td>-0.0084</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.18)</td>
<td>(0.19)</td>
<td>(0.045)</td>
<td>(0.097)</td>
<td>(0.089)</td>
<td>(0.047)</td>
<td>(0.047)</td>
<td>(0.028)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.84</td>
<td>0.18</td>
<td>0.32</td>
<td>0.46</td>
<td>0.26</td>
<td>0.35</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>Mother’s ed. missing</td>
<td>RD estimate</td>
<td>0.0018</td>
<td>0.0086</td>
<td>-0.0041</td>
<td>0.0024</td>
<td>-0.0086</td>
<td>-0.0040</td>
<td>-0.00042</td>
<td>-0.0034*</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.0080)</td>
<td>(0.0094)</td>
<td>(0.0028)</td>
<td>(0.0033)</td>
<td>(0.0066)</td>
<td>(0.0034)</td>
<td>(0.0027)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.82</td>
<td>0.40</td>
<td>0.20</td>
<td>0.61</td>
<td>0.17</td>
<td>0.39</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Father’s ed. missing</td>
<td>RD estimate</td>
<td>-0.019*</td>
<td>-0.0081</td>
<td>-0.0037</td>
<td>-0.0014</td>
<td>-0.0046</td>
<td>-0.0052</td>
<td>-0.0071***</td>
<td>-0.0046**</td>
</tr>
<tr>
<td></td>
<td>conv. s.e.</td>
<td>(0.0094)</td>
<td>(0.0096)</td>
<td>(0.0026)</td>
<td>(0.0043)</td>
<td>(0.0063)</td>
<td>(0.0037)</td>
<td>(0.0027)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td></td>
<td>robust p</td>
<td>0.047</td>
<td>0.36</td>
<td>0.19</td>
<td>0.71</td>
<td>0.44</td>
<td>0.32</td>
<td>0.011</td>
<td></td>
</tr>
</tbody>
</table>

Note: Robust semi parametric sharp RD estimates of the effect of paternity leave quotas on balancing variables 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 weighted average of reform-specific estimates. Robust, but not bias-corrected standard errors reported. "robust p"-values are bias-corrected. * p<0.1 ** p<0.05 *** p<0.01, based on the robust, but not bias-corrected standard errors (themselves not reported).
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.
A.3 Additional Results

Figure A3 reports the impacts on mothers’ and fathers’ annual incomes over time using the stacked and pooled fuzzy RD estimates. 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 we can rule out positive impacts larger than around NOK 5,000 on mother’s annual earnings in response to each week of paternity leave use for all years post-birth.

Figure A3: Fuzzy and stacked RD estimates of the effects of paternity leave use on mothers’ and fathers’ earnings.

Note: Left figure shows fuzzy RD estimates of the impact of an additional week of paternity leave use on mother’s earnings over time, using all six reforms. Right figure shows fuzzy RD estimates of the impact of paternity leave use on father’s earnings over time, where confidence intervals for the pooled specification has been capped at +20,000 NOK to maintain readability of the axis. Pooled estimate refers to the weighted average of reform-specific estimates, while the stacked estimate stacks the cutoff-specific specifications. Robust bias-correcting inference reported for the pooled estimate, conventional, heteroskedasticity-robust inference for the stacked estimate.

In Figure A4 we report estimates for each reform separately. This involves estimating the following equations:

\[ y_{it} = \beta_t L_i + f_t(x_i) + \mu_{it} \]
\[ L_i = \gamma \mathbb{1}(x_i \geq 0) + g(x_i) + \xi_{it} \]

(5)
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. 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 in order to obtain estimates local to the cutoff. We estimate and report robust bias-corrected confidence intervals (Calonico et al. 2014) together with the conventional, heteroskedasticity-robust confidence intervals. We then scale the effects on earnings to reflect the percentage changes in the child penalty. Many models in this section are estimated using the robust RD commands for Stata written by Matias D. Cattaneo et al., to whom we owe thanks. These include rdrobust, rddensity and rdbwselect. See Calonico et al. (2018) and Cattaneo et al. (2018).
Figure A4: Robust RDD estimates, paternity leave reforms

Notes: First and third columns show binned plots of the weeks of paternity leave against birth date of the child in days after the reform, overlaid with the estimated local linear polynomials. Second and fourth panels show sharp RD estimates of the impact of an additional week of quota on maternal and paternal income by year. Optimal MSE-reducing bandwidths, triangular kernel and local linear polynomials on either side of cutoff. Confidence intervals are robust and bias-corrected.
Table A3: Paternity norms: Fuzzy RD of paternity leave on leave for next child for all years separately

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RD estimate per week</td>
<td>-0.313</td>
<td>-0.501</td>
<td>-0.0624</td>
<td>-0.256</td>
<td>0.79**</td>
<td>-0.268</td>
<td>-0.092</td>
</tr>
<tr>
<td>conv. standard error</td>
<td>(1.08)</td>
<td>(0.499)</td>
<td>(0.123)</td>
<td>(0.314)</td>
<td>(0.359)</td>
<td>(0.359)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>robust standard error</td>
<td>1.32</td>
<td>0.592</td>
<td>0.147</td>
<td>0.366</td>
<td>0.404</td>
<td>0.412</td>
<td></td>
</tr>
<tr>
<td>conventional p-value</td>
<td>0.772</td>
<td>0.315</td>
<td>0.612</td>
<td>0.416</td>
<td>0.003</td>
<td>0.428</td>
<td>0.40</td>
</tr>
<tr>
<td>robust p-value</td>
<td>0.873</td>
<td>0.284</td>
<td>0.605</td>
<td>0.403</td>
<td>0.024</td>
<td>0.491</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>14,201</td>
<td>14,761</td>
<td>14,086</td>
<td>9,704</td>
<td>704</td>
<td>53,456</td>
<td>53,456</td>
</tr>
<tr>
<td>Optimal bandwidth</td>
<td>60.9</td>
<td>48.4</td>
<td>60.4</td>
<td>46.8</td>
<td>46.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficient observations</td>
<td>4,821</td>
<td>4,245</td>
<td>4,893</td>
<td>2,778</td>
<td>159</td>
<td>16,896</td>
<td>16,896</td>
</tr>
</tbody>
</table>

Notes: Fuzzy RD estimates of the impact of one more week of paternity leave for a child on the weeks of paternity leave use for the next. Conventional standard errors are heteroskedasticity-robust but not bias corrected. **p<0.01, *p<0.05, *p<0.1 based conventional standard errors.

A.4 Accounting for Effects of Maternal Leave

As evident from Table 2a, several of the reforms affected not only the paternity leave quota, but also the maternity leave quota and the sum of the maternity leave quota and the shared leave. As documented in Table A2, this resulted in reduced maternity leave take-up roughly for the reforms where the total time a mother could take off work was reduced. Although we argue that this change in maternity leave takeup is relatively minor compared with 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:

\[
\begin{align*}
  y_{ir_t} &= \beta^L_i L_{i_t} + \beta^M_i M_{i_t} + \varphi^0_i x_{i_t} \mathbb{1}(x_{i_t} < 0) + \varphi^1_i x_{i_t} \mathbb{1}(x_{i_t} \geq 0) + \pi_r + \epsilon_{ir_t} \\
  L_{i_r} &= \gamma^Q_i Q_{i_r} + \gamma^S_i S_{i_r} + \varphi_i^L x_{i_r} \mathbb{1}(x_{i_r} < 0) + \varphi_i^L x_{i_r} \mathbb{1}(x_{i_r} \geq 0) + \pi^L_r + \eta^L_{ir} \\
  M_{i_r} &= \gamma^MQ_i Q_{i_r} + \gamma^MS_i S_{i_r} + \varphi_r^M x_{i_r} \mathbb{1}(x_{i_r} < 0) + \varphi_r^M x_{i_r} \mathbb{1}(x_{i_r} \geq 0) + \pi^M_r + \eta^M_{ir} 
\end{align*}
\]
where $L_{ir}$ and $M_{ir}$ are paternity and maternity leave takeup for couple $i$ who is exposed to reform $r$. Rather than a dummy at the cutoff, the instruments are now $Q_{ir}$, the paternity leave quota, and $S_{ir}$, the sum of shared leave and maternity leave quota. Notice that the variation in these 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 for each reform and a triangular kernel to control for the forcing variable. The outcome variable $y_{irt}$ is labor market earnings, measured separately for mothers and fathers. This leaves us with two treatments by two outcomes per year we measure outcomes.

When instrumenting for two endogenous variables in an IV-setup, it is not clear how to determine the optimal MSE-reducing bandwidth as before. We therefore use a) the MSE-reducing optimal bandwidth for the first stage of either of the instruments or b) a fixed 50-day bandwidth. As before, we report robust, but not bias-corrected standard errors for the stacked specification.

First stage results for the two endogenous variables are reported in Table A4. Notice that independent variation to identify both effects relies on stacking all reforms, so that we cannot perform these estimates separately by reform. The choice of bandwidth is not of essence: The results are very similar whether we use either the MSE-reducing optimal bandwidths or a fixed 50-day window. Second, note that the reforms work exactly as we would expect: An increase in the daddy quota of 1 week increases paternity leave uptake by almost exactly 1 week when we control for changes to the remaining quota for the mother. 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 A4: First stage effects of maternity and paternity leave quotas

<table>
<thead>
<tr>
<th>Weeks of leave</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mother</td>
</tr>
<tr>
<td>A: 50-day bandwidth</td>
<td></td>
</tr>
<tr>
<td>Paternity leave quota $(Q_{pr})$</td>
<td>0.066</td>
</tr>
<tr>
<td>(0.14)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Remaining leave for mother $(S_{pr})$</td>
<td>0.77***</td>
</tr>
<tr>
<td>(0.21)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>joint $F$</td>
<td>21.1</td>
</tr>
<tr>
<td>$N$</td>
<td>24,520</td>
</tr>
<tr>
<td>B: Maternity leave-optimal bandwidth</td>
<td></td>
</tr>
<tr>
<td>Paternity leave quota $(Q_{pr})$</td>
<td>0.069</td>
</tr>
<tr>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Remaining leave for mother $(S_{pr})$</td>
<td>0.79***</td>
</tr>
<tr>
<td>(0.20)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>joint $F$</td>
<td>24.8</td>
</tr>
<tr>
<td>$N$</td>
<td>29,028</td>
</tr>
<tr>
<td>C: Paternity leave-optimal bandwidth</td>
<td></td>
</tr>
<tr>
<td>Paternity leave quota $(Q_{pr})$</td>
<td>0.055</td>
</tr>
<tr>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Remaining leave for mother $(S_{pr})$</td>
<td>0.73***</td>
</tr>
<tr>
<td>(0.21)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>joint $F$</td>
<td>18.0</td>
</tr>
<tr>
<td>$N$</td>
<td>27,344</td>
</tr>
</tbody>
</table>

Note: 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 eq. 6. Panel A) uses a fixed 50-day bandwidth, panel B) uses the MSE-reducing optimal bandwidth for each reform if instrumenting for maternity leave only, panel C) the same for paternity leave. Heteroskedasticity robust, but not bias-corrected standard errors. *$p<0.1$, **$p<0.05$, ***$p<0.01$

Because the choice of bandwidth does not seem to matter and because we’re primarily interested in the effects of paternity leave, we present fuzzy stacked RD estimates based on this specification using the paternity leave-optimal bandwidth from panel C. As in the base model in the paper we also revert to the reform-specific dummies as instruments when reporting the IV estimates rather than quota measures.

Results from the stacked fuzzy RD model where we instrument for both mothers’ and fathers’ leave take up is presented in Figure A5. 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. Except perhaps for the outlier at child age 4, 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 A5: Effects of maternity and 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
(c) Maternity leave on mother’s earnings
(d) Maternity leave on father’s earnings

Note: Top panels 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 eq. For comparison we also show our stacked fuzzy RD estimates from the baseline model where we only instrument for the weeks of paternity leave. Bottom panels show the impact of an additional week of maternity leave on mothers’ (left) and fathers’ (right) earnings.

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  Child Care: Robustness and Additional Results

Figure B1: Predicting expansion of slots from pre-reform characteristics

(a) Child care coverage for 1-year olds

(b) Child care coverage for 2-year olds

Note: Results from regression of our two instruments, child care 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 shows the impact of high equality, subsidized early child care on each parent’s earnings. Focusing first on the years of treatment, ages 1-3, we see that the estimates increase in this period up to point estimates of around 27,000 NOK at age 2 and close to 30,000 NOK at age 3, where most of the treatment happens, only to return to zero the last two years of the panel. Estimates are significant at the 5 percent level at age 3 and 10 percent level at age 2, and thus indicate that there is some immediate effects of use on earnings during the years of treatment, perhaps driven by allowing mothers to return to work earlier after child birth. Results for fathers are noisy, but point, if anything, to negative impacts on earnings, which could also reduce child penalties. The pre-birth outcomes, which we can think of as placebo outcomes, indicate small and insignificant impacts of future child care use on past earnings, supporting the estimation strategy.

---

27 This estimate is smaller than the baseline estimate in Andresen and Havnes (2019), but a number of differences in the sample and specification may explain this, as well as the lower level of precision in our study due to a sample size than half the size because of the focus on first born children only.
Figure B2: Impact of early child care use on mother’s and father’s earnings

(a) Mothers earnings, unscaled
(b) Fathers’ earnings, unscaled

Note: IV results from equation 4 reflecting the impact on labor earnings in 1,000 NOK across child age for an extra year of early child care use at ages 13-36 months on mothers’ and fathers’ earnings, from eq. 4 in the main paper.