

NBER WORKING PAPER SERIES

DO FAMILY POLICIES REDUCE GENDER INEQUALITY? EVIDENCE FROM
60 YEARS OF POLICY EXPERIMENTATION

Henrik Kleven
Camille Landais
Johanna Posch
Andreas Steinhauer
Josef Zweimüller

Working Paper 28082
<http://www.nber.org/papers/w28082>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020, Revised January 2021

We thank Francine Blau, Hilary Hoynes, Emmanuel Saez, Johannes Spinnewijn, Gabriel Zucman, and seminar participants at the AEA Meetings 2019, Bank of Italy, UC Berkeley, Bocconi, Crest-Ecole Polytechnique, EIEF, OeGB Innsbruck, PSE, National Bank of Belgium, University of Reading, Stockholm University, University of Salzburg, Tinbergen Institute and Queen Mary. We gratefully acknowledge support from the Center for Economic Behavior and Inequality (CEBI) at the University of Copenhagen, financed by grant #DNR134 from the Danish National Research Foundation. Camille Landais gratefully acknowledges support from the LSE International Inequality Institute and from ERC grant #679704-DYNAMICSS. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation
Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller
NBER Working Paper No. 28082
November 2020, Revised January 2021
JEL No. H31,H42,J08,J13,J16,J18,J22

ABSTRACT

Do family policies reduce gender inequality in the labor market? We contribute to this debate by investigating the joint impact of parental leave and child care, using administrative data covering the labor market and birth histories of Austrian workers over more than half a century. We start by quasi-experimentally identifying the causal effects of all family policy reforms since the 1950s on the full dynamics of male and female earnings. We then map these causal estimates into a decomposition framework building on Kleven, Landais and Søgaard (2019) to compute counterfactual gender inequality series. Our results show that the enormous expansions of parental leave and child care subsidies have had virtually no impact on gender convergence.

Henrik Kleven
Department of Economics
Princeton University
238 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and CEPR
and also NBER
kleven@princeton.edu

Andreas Steinhauer
University of Edinburgh
31 Buccleuch Place
Edinburgh EH8 9JT
United Kingdom
and CEPR
andreas.steinhauer@ed.ac.uk

Camille Landais
Department of Economics
London School of Economics
Houghton Street
London WC2A 2AE
United Kingdom
and CEPR
c.landais@lse.ac.uk

Josef Zweimüller
Department of Economics
University of Zurich
Schoenberggasse 1
8001, Zurich
Switzerland
josef.zweimueller@econ.uzh.ch

Johanna Posch
Analysis Group
1 Angel Court
London
EC2R 7HJ
UK
Johanna.Posch@analysisgroup.com

1 Introduction

Is government intervention critical for reducing gender inequality in the labor market? Or is gender inequality driven primarily by factors outside the government's control? Recent research shows that the bulk of gender inequality can be attributed to the unequal impacts of parenthood on men and women (Bertrand, Goldin and Katz 2010; Kleven, Landais and Sogaard 2019; Kleven, Landais, Posch, Steinhauer and Zweimuller 2019), which suggests that family policies such as parental leave and child care have the greatest potential to affect gender inequality. A sizeable literature studies the impact of family policies on female labor market outcomes, typically focusing on the contemporaneous impacts of marginal reforms in isolation. Apart from cross-country evidence, there is little research trying to estimate the aggregate long-run impact of public policy on gender inequality (Olivetti and Petrongolo 2017). To make progress on this question, we investigate the joint impact of parental leave and child care over more than half a century. Our main conclusion is simple and striking: The enormous expansions of parental leave and child care subsidies have had virtually no impact on gender convergence.

To reach this conclusion, we rely on exceptional administrative data covering the labor market and birth histories of Austrian workers from 1953 to 2017. During this era, the Austrian gender gap in earnings fell by about 30 percentage points, while public policies supporting families with children were introduced and greatly expanded. Parental leave was introduced in 1961 and subsequently expanded (and in one instance scaled back) through a series of major reforms. At the same time, the provision of heavily subsidized public child care was rolled out across the country. Similar expansions of family policies have been implemented in other European countries and their implications for gender inequality is the subject of heated debate: Is access to job-protected parental leave and subsidized child care instrumental for boosting female labor force participation and reducing gender gaps? Are generous parental leave schemes counterproductive by inducing mothers to stay at home rather than investing in their careers? Or is the emphasis on policy solutions overstated, with gender inequality being driven mostly by equilibrium features of the labor market such as the temporal flexibility of jobs (Goldin 2014; Goldin and Katz 2016) and gender norms or culture (Bertrand 2011; Kleven and Landais 2017; Kleven, Landais, Posch, Steinhauer and Zweimuller 2019; Boelmann, Raute and Schönberg 2020)?¹

¹Another hypothesis is that gender gaps are endemic to biological gender differences, but Kleven, Landais and Sogaard (2020) provide evidence against the importance of such channels.

We take advantage of the richness of the Austrian administrative registers combined with quasi-experimental techniques to implement a unique, bottom-up approach to measuring counterfactual gender inequality absent family policies. Our approach proceeds in two steps. First, we identify the causal effects of *all* family policy reforms since the 1950s on the labor market outcomes of men and women. In a second step, we map these causal estimates into a decomposition framework to compute counterfactual gender gaps absent policy intervention.

Besides our long-run historical approach, we add two key innovations to the existing literature. The first innovation is to use “child penalties” in earnings as our main outcome. This ensures that we identify the full dynamic effect of each policy change on the labor market trajectories of women relative to men. The second innovation is to account for the potential interaction effects between parental leave and child care policies. Family policies are often designed as a bundle, with child care coverage starting around the time that parental leave eligibility is ending. As a result, we analyze the effects of parental leave and child care policies jointly, capturing any non-linearities brought about by their potential interaction.

Leveraging a regression discontinuity design based on cutoff dates of the parental leave reforms, we first show that parental leave duration has significant negative effects on the earnings of women relative to men in the short run. We show that these effects are non-linear, being greater at lower baseline durations. However, for all parental leave reforms, we find precisely estimated zero effects in the long run. Leveraging a difference-in-differences design similar to [Duflo \(2001\)](#) based on large expansions in locally provided child care, we find precisely estimated zero effects of both nursery care and pre-school care on the earnings of women relative to men.² We provide suggestive evidence that this null effect is driven by crowding out of other types of informal care, and the persistence of strong norms for maternal care. Finally, we show that there are no interaction effects between child care and parental leave policies.³

In the second step of our analysis, we use these estimates to back out counterfactual long-run series of gender inequality. We build on the framework developed by [Kleven, Landais and Sogaard \(2019\)](#), which decomposes gender inequality using cohort-specific estimates of child penalties. The intuition is simple: because we know how each policy reform causally impacted child penalties, we can reconstruct cohort-specific child penalties (and therefore aggregate gender inequality over time) in the absence of policy intervention. Our results imply that the dramatic expansion of family policies has had

²Similar designs using local supply shocks have been used to study the impact of child care on female employment (e.g, [Havnes and Mogstad 2011](#)) and children’s outcomes (e.g., [Cornelissen, Dustmann, Raute and Schönberg 2018](#)).

³The absence of any interaction effects between child care provision and parental leave is relatively unsurprising *ex post* (but not *ex ante*) given our finding that the effects of the two policies separately are zero.

a remarkably small impact on gender inequality. Due to the negative effects of parental leave on labor supply in the short-run, family policies have been slightly counterproductive for gender convergence. The average earnings of Austrian women is 43.5% lower than the average earnings of men today, while they would have been 41.5% lower if family policies had remained as they were in 1960. Sensitivity analyses suggest that this null effect is very precise. We also demonstrate that our conclusions are robust to accounting for frictions, equilibrium effects, and fertility responses.⁴

Our paper contributes to the large literature on gender inequality in the labor market (reviewed by [Bertrand 2011](#) and [Olivetti and Petrongolo 2016](#)) and especially to recent work showing the crucial role of parenthood ([Bertrand, Goldin and Katz 2010](#); [Kleven, Landais and Sogaard 2019](#); [Kleven, Landais, Posch, Steinhauer and Zweimuller 2019](#); [Kuziemko, Pan, Shen and Washington 2018](#); [Cortés and Pan 2020](#)). Our paper also contributes to a burgeoning literature on the impact of family policies (reviewed by [Olivetti and Petrongolo 2017](#)), including parental leave schemes (see e.g., [Lalive and Zweimüller 2009](#); [Schönberg and Ludsteck 2014](#); [Dahl, Løken, Mogstad and Salvanes 2016](#)) and child care subsidies (see e.g., [Baker, Gruber and Milligan 2008](#); [Havnes and Mogstad 2011](#)). We discuss how our findings relate to these literatures in more detail below. While we are far from the first to study these questions, our paper sheds new light by studying them within a unified quasi-experimental setting in order to quantify the aggregate importance of family policies for the long-run historical evolution of gender inequality.

Our bottom-up counterfactual methodology offers a credible approach to measuring the impact of policy interventions on the long-run evolution of gender inequality. The conclusions stand in sharp contrast to those that can be gleaned from cross-country approaches, which tend to find strong correlations between public policy and female labor market outcomes (e.g., [Ruhm 1998](#); [Blau and Kahn 2013](#); [Kleven 2014](#); [Olivetti and Petrongolo 2017](#)). The conceptual approach developed here could be used to study the role of policy intervention in other areas of inequality research (such as income and wealth inequality).

The remainder of the paper is organized as follows. Section 2 describes the Austrian institutional context and data. Section 3 presents quasi-experimental evidence on the impact of parental leave reforms on child penalties. Section 4 presents quasi-experimental evidence on the impact of local child care expansions on child penalties. This section also provides evidence on the potential mechanisms behind our null effect. Section 5

⁴Our null result is striking given that the subsidies to parental leave and child care in Austria (and in other European countries) create large changes in participation tax rates (accounting for all taxes, transfers, and subsidies) during the years following child birth. In fact, the implications of these policies for the extensive margin incentives of women with young children tend to be much larger than the implications of, for example, the famed EITC program in the U.S. See [Kleven \(2014\)](#) for a cross-country analysis of participation tax rates and labor supply.

lays out our decomposition framework, and provides estimates of the counterfactual evolution of gender inequality under different policy trajectories. Section 6 concludes.

2 Context and Data

2.1 Institutional Context

A (Typical) Gender-Conservative Environment Austria, like other German-speaking countries, is characterized by relatively large differences in the labor market outcomes of men and women. The raw gender gap in earnings is roughly 40% and the female labor force participation rate is about 55% (in 2018). However, it is worth highlighting two factors that make Austria an interesting laboratory for studying the impact of family policies on gender inequality. First, Austria is not an international outlier in terms of gender gaps. English-speaking countries like the U.S. and the U.K. have broadly similar levels of gender inequality in earnings. Moreover, female labor force participation in Austria is higher than in a number of other European countries such as Italy, Spain, and France. Second, Austria has experienced a substantial reduction in gender gaps and a surge in female labor force participation over the past 50 years. This makes it a very useful setting to study whether this reduction in gender inequality is causally related to family policy expansions.

German-speaking countries are often singled out for having relatively conservative gender norms. Appendix Figure A.I probes the idea that gender norms are more conservative in Austria than elsewhere. The evidence presented is based on questions from the 2012 wave of the International Social Survey Program (ISSP) on the attitudes towards market work by women with and without children. Two insights emerge from the figure. One is that gender attitudes are quite traditional in most countries: large fractions of people hold the view that women with young children should stay at home rather than work in the market (panel (a)), and that, if they don't, their children will suffer (panel (b)). The other insight is that Austrians, while being on the more conservative side of the spectrum, are not so different from the citizens of other countries. Their views on whether mothers with young children should stay at home are very similar to the views observed in the U.S., the U.K., and France. Overall, the evidence in Figure A.I suggests that Austria is gender conservative, but not a strong outlier internationally. Austria is a typical gender-conservative environment, and as such, a great setting for our study. We are interested in better understanding whether family policies can change the evolution of gender inequality when gender attitudes are still traditional.

Family Policies Over the last 60 years, most developed countries have introduced and dramatically expanded two sets of family policies: parental leave and child care provision (Olivetti and Petrongolo 2017). Austria is a case in point.

In the aftermath of World War II, there was no system of parental leave in Austria. Women were entitled to *Mutterschutz*, a maternity insurance period of 8 weeks following delivery, akin to sick leave. In 1961, a paid and job-protected maternity leave of 52 weeks was introduced. Since then, the country has implemented four large reforms of its parental leave policies (in 1990, 1996, 2000, and 2008), which have considerably expanded the generosity and scope of parental leave.⁵ Today, the Austrian parental leave system stands out as one of the most generous in the developed world. Currently, parents are entitled to up to about 35 months of paid parental leave if that time is shared, or up to about 28 months if one parent takes the time alone. The replacement rate is relatively high: the median replacement rate is about 40% of net earnings. The take-up of parental leave comes with a job-protection guarantee of 24 months. The receipt of parental leave benefits is not subject to any work eligibility requirement prior to child birth.

Institutional child care in Austria is divided into nursery care (*Krippe*) covering children from ages 1 to 2 and pre-school care (*Kindergarten*) from ages 3 to 5. The provision of both forms of institutional child care has been greatly expanded over the past 50 years. In the beginning of the 1970s, less than 2% of children between the age of 1 and 2 attended nursery. Today, this number has risen to 25%. The fraction of children attending pre-school has risen from 40% to more than 90% over the same time period. Child care is provided by municipalities, following general guidelines set by state legislation.⁶ The median number of daily operating hours for Austrian child care institutions is 9 hours, with typical opening hours from 7:30am to 4:30pm.⁷ Institutional child care has always been heavily subsidized in Austria: out-of-pocket costs are among the lowest in developed nations. According to the OECD Family Database, out-of-pocket costs for a family with two children aged 2 and 3 in full-time daycare was just 3% of family net income in 2015, compared to an average of 13% across the OECD and 22% in the U.S.⁸ Note that subsidized institutional child care encompasses both public and church provided care. The private market for child care is very limited and essentially consists of a few private creches and child-minders (*Tagesmutter*).⁹

⁵Two further reforms in 2010 and 2017 provided additional flexibility to parents regarding the length of leave and the amount of the transfer. We do not evaluate these policy reforms here because the time-horizon to study future earnings dynamics is too short.

⁶State legislation typically determines the opening hours, the fee structure, and the total number of children that can be supervised by a teacher. For details see Baierl and Kaindl (2011).

⁷See Baierl and Kaindl (2011).

⁸<http://www.oecd.org/els/family/database.htm>. Note that small variations in prices exist across states, and sometimes across municipalities.

⁹Figures from a microcensus survey in 2002 reveal that less than 3% of children under age two were

2.2 Data

Labor Market Outcomes & Birth History Our study relies on unique data covering the labor market and birth histories of Austrian workers from 1953 to 2017. Our main data come from the Austrian Social Security Database (ASSD, see [Zweimüller, Winter-Ebmer, Lalive, Kuhn, Wuellrich, Ruf and Büchi 2009](#) for details). This is a matched employer-employee dataset, which contains detailed information on employment and earnings by calendar year from 1972 to 2017. The ASSD register covers the universe of paid employment spells, with the notable exclusion of civil servants, self-employed and farmers. This means that about 85% of the Austrian population is covered by ASSD, and this coverage rate has been stable over time.

Earnings recorded in the ASSD correspond to the amount on which workers pay social security contributions, which is subject to a yearly-adjusted cap. As a result, our earnings variable is top-coded for about 5% of workers. To deal with top-coding, we replace all right-censored earnings with mean earnings above the cap predicted from a Pareto distribution, the parameters of which we obtain by gender and age groups using uncensored income tax data.¹⁰ Besides earnings, the ASSD includes information on the uptake of all social insurance programs along with demographic variables such as gender, year of birth, year of death, nationality, and all child birth events of women (e.g., date of birth of each child).

ASSD data exist before 1972, but has not been digitized. However, detailed tabulations of earnings by gender, drawn from the original data, have been regularly published. This enables us to consistently compute the evolution of gender inequality on the population covered by the ASSD since 1953. Furthermore, we use additional micro-level earnings information before 1972 from the *Rueckwirkend Erfasste Versicherungsdaten* (REV) register. The REV retrospectively digitized, for ASSD workers retiring in the 1980s and 1990s, earnings and employment spells prior to 1972 for the purposes of pension calculations.¹¹ The selective nature of this data means that only certain cohorts are well covered. Fortunately, for birth cohorts 1920-1950, which are primarily those having children between 1955-1972, the REV covers about 80% of workers, as shown in Appendix Figure D.V. Although the REV does not contain direct information on child births, it has information on the uptake of maternity insurance (*Mutterschutz*), which has been available to Austrian women ever since the 1950s. As a consequence, the REV data enables us to consistently estimate child penalties for women giving birth all the way back to the mid-1950s. We are not aware of any existing work estimating child penalties over such a long time period.

taken care of by *Tagesmutter*.

¹⁰We provide further details on our methodology and its robustness in Appendix section D.1.

¹¹The earnings information is provided in the same way as in the ASSD (except there is no information about the firm), and we adjust for topcoding using the same Pareto estimates.

Importantly, both the ASSD and REV contain child birth information only for women. We can nevertheless link children to their fathers for a large subset of child births. To this end, we use the child benefit register (*Beihilfedaten*). This register contains the universe of child benefit claims filed between 1995-2012. As the take-up of child benefits is almost complete, we are able to match around 90% of children born between 1995 and 2012 cohorts. Each claim lists the social security number of the claimant (usually the mother) as well as that of her partner. We pick the first entry in a child’s life and set the man listed on that claim as the father.

Child Care Provision Our analysis also relies on granular information on child care provision at the municipal level. We have obtained administrative data collected by Statistics Austria, which records, for every year since 1988 and for each of the nearly 2,000 Austrian municipalities, the number of child care facilities (broken down by nurseries and pre-schools), their opening hours (full-day or half-day), and the number of teachers in each facility. From this data, we are able to construct precise measures of child care supply at the municipal level. We provide details about the construction of these measures in section 4.

2.3 Child Penalties

Our analysis of the impact of family policies on gender inequality over the last 60 years relies on two inputs. First, we need to estimate how the full dynamics of parental outcomes respond to the arrival of children, what we refer to as *child penalties*. Second, we need to identify the causal effect of all family policy reforms on these child penalties. In this section, we describe how child penalties are defined and estimated, and document their magnitude in the Austrian context.

To estimate the impact of children on the labor market trajectories of mothers and fathers, we adopt the event-study specification proposed by [Kleven, Landais and Sogaard \(2019\)](#). For each parent in the data, event time t is indexed relative to the year of the first child birth. Denoting by Y_{ist}^g the outcome for individual i of gender g in year s and at event time t , we run the following regression separately for men and women:

$$Y_{ist}^g = \boldsymbol{\alpha}^g \mathbf{D}_{ist}^{Event} + \boldsymbol{\beta}^g \mathbf{D}_{ist}^{Age} + \boldsymbol{\gamma}^g \mathbf{D}_{ist}^{Year} + v_{ist}^g \quad (1)$$

On the right-hand side, we use boldface to denote vectors. The first term includes event time dummies, indexed such that $t = 0$ denotes the year of arrival of the first child.¹²

¹²For notational convenience, we get rid of the transpose operator and denote by $\boldsymbol{\alpha} \mathbf{D}$ the scalar product of two vectors of same dimension, $\boldsymbol{\alpha}$ and \mathbf{D} .

We omit the event-time dummy at $t = -1$, implying that the event-time coefficients measure the impact of children relative to the year just before the first child. The second term includes age dummies (to control for lifecycle trends) and the third term includes year dummies (to control for time trends). We are able to identify the effects of all three sets of dummies because, conditional on age and year, there is variation in event time driven by variation in the age at which individuals have their first child.¹³

Our main outcome variable is gross labor earnings, specified in levels.¹⁴ We convert the estimated level effects into percentages to obtain the child penalty P_t^g for gender g at event time t :

$$P_t^g \equiv \frac{\hat{\alpha}_t^g}{E[\tilde{Y}_{ist}^g | t]} \quad (2)$$

where \tilde{Y}_{ist}^g is the predicted outcome when omitting the contribution of the event time dummies. This implies that P_t^g can be interpreted as the percentage loss of average earnings due to having children.¹⁵

Figure 1 shows estimated child penalties for different outcomes in our ASSD sample. The magnitude of earnings penalties on mothers is very large (panel (a)). Ten years after the birth of the first child, Austrian mothers suffer a 50% earnings drop relative to fathers. Panels (b), (c) and (d) show that this effect reflects a combination of penalties along three margins: the extensive margin of labor supply, the intensive margin of labor supply, and the wage rate. Although these patterns are qualitatively similar to those documented in Scandinavia, the magnitudes are significantly larger in Austria. As shown in Appendix Figure A.II, long-run earnings penalties in Scandinavia lie in the 20-30% range, about half the size of the long-run penalties observed in Austria and Germany.

On average, over our period of interest, Austrian mothers have thus faced large career costs from having children. In the next two sections, we turn to analyzing how family policies have affected the evolution of these penalties.

¹³Kleven, Landais and Sogaard (2019) lay out the identification assumptions underlying this approach, compare its results to alternative approaches in the literature, and provide evidence of its ability to identify the causal effect of parenthood.

¹⁴We specify equation (1) in levels rather than in logs to be able to keep the zeros in the data due to nonparticipation.

¹⁵Importantly, under the identifying assumptions, equation (2) corresponds to $\frac{\hat{\alpha}_t^g}{E[\tilde{Y}_{ist}^g | t]} = \frac{E[y_{1,t}|t] - E[y_{0,t}|t]}{E[y_{0,t}|t]}$, where $y_{1,t}$ is the actual earnings at event time t and $y_{0,t}$ is the counterfactual earnings absent children. In contrast, a log specification would estimate the average percentage earnings loss, which differ if treatment effects are correlated with counterfactual earnings. That is, if for example the effect of children on earnings is higher or lower for women who would have had higher earnings absent children. P_t^g is therefore the relevant metric when assessing the effect of children on the aggregate gender pay gap.

3 The Impact of Parental Leave on Child Penalties

As mentioned above, Austria has considerably expanded the scope and generosity of paid parental leave (PL). We now investigate how the dramatic expansion of parental leave since the 1960s has affected the labor market outcomes of Austrian mothers and fathers.

As a starting point, we focus on the parental leave regime in place since 2008. This regime offers parents the possibility of choosing between three PL options: (i) a long parental leave (“30+6”), with 30 months for the mother and 6 months for the father at a daily flat benefit of €14.53; (ii) an intermediate parental leave (“20+4”), with 20 months for the mother and 4 months for the father at a daily benefit of €20.80; and (iii) a short parental leave (“15+3”), with 15 months for the mother and 3 months for the father at a daily rate of €26.60.¹⁶ In Figure 2, we split parents of children born in 2008 into three groups, according to the PL option they selected into, and report the child penalty P_t^g for each group separately. The figure shows that the earnings trajectory of fathers is unaffected by children irrespective of the selected PL option. By contrast, the figure shows stark differences in child penalties across PL options for mothers. The magnitude and persistence of the earnings drop after child birth correlates strongly with the length of their PL selection. Nine years after child birth, women in the “30+6” scheme exhibit penalties that are about 15pp larger than the penalties of women in the “15+3” scheme.

The strong correlation between PL duration and female child penalties begs the question of whether it is causal. In other words, how much of the correlation reflects a causal effect of PL duration, and how much reflects differential selection into PL options? To answer this, we turn to the four large reforms that exogenously changed the duration of parental leave over the past 60 years.

3.1 The 1990, 1996, and 2000 Reforms

Empirical Strategy We start by focusing on the three reforms that took place in 1990, 1996, and 2000. These reforms share important institutional similarities that allow us to study them using the same empirical strategy.

First, all three reforms focused on the *duration* of the paid parental leave. The 1990 reform increased the maximum duration of leave from 12 to 24 months. The 1996 reform introduced the rule that each parent had to take at least 6 months of leave, thus

¹⁶Households need to choose one of these three options in the 8 weeks following the birth of the child. Their choice is legally binding and cannot be altered ex-post.

reducing maximum *maternal* leave from 24 to 18 months. Finally, the 2000 reform increased the maximum duration of maternal leave from 18 to 30 months.¹⁷ Importantly, the other parameters of the PL system did not change after 1990, including job protection provisions and benefit levels. As for job protection, since 1990, a woman has had the right to return to her previous employer for up to 24 months after child birth. As for benefits, until 2008, they were a flat amount (indexed to wage inflation) that did not depend on previous earnings or the duration of leave.¹⁸ As a result, the 1996 and 2000 reforms allow for identifying the effects of changing PL duration, conditional on all other PL parameters, while the 1990 reform allows for identifying the effect of changing the duration of both paid leave and job protection.

Second, a key feature of all three reforms is that eligibility was based on a cutoff date for the birth of the child, with no grandfathering. For the 1990 reform, mothers of children born after July 1st 1990 became eligible to the new scheme, offering a maternity leave entitlement of 24 months. Mothers of children born before this date remained entitled to just 12 months of leave. The corresponding cutoff dates for the 1996 and 2000 reforms were July 1st 1996 and July 1st 2000. This implementation structure lends itself to a regression discontinuity design, using the date of birth of the first child as the assignment variable.¹⁹

Regression Discontinuity Evidence We are interested in the effect of parental leave reforms on the full dynamics of female outcomes around the event of the first child birth.²⁰ To study this effect, Appendix Figure B.I starts by displaying, for all event years $t \in [-4; 10]$, the average earnings of mothers in weekly bins of the running variable around the cutoff date of the 1990 reform. Starting with panel (a), we see that earnings

¹⁷Before 1990, the parental leave benefits were only available to mothers. The 1990 reform offered the possibility for couples to split the total credit of 24 months between the mother and the father. Fathers were then entitled to take up to 6 months out of the 24 months, but could only take their leave in the second year of the child. In practice, the take up rate of paternal leave was zero. We therefore treat the 1990 reform as an increase in the potential duration of maternal leave from 12 to 24 months. The 1996 reform kept the total duration of parental leave to 24 months, of which one parent could not take more than 18 months. We therefore treat the 1996 reform as a decrease in the potential duration of maternal leave from 24 to 18 months. The 2000 increased the maximum duration of parental leave to 36 months, of which one parent could not take more than 30 months. We therefore treat the 2000 reform as an increase in the potential duration of maternal leave from 24 to 18 months.

¹⁸Note also that prior work history requirements did not change over the period 1990 to 2000, stipulating that an individual needed to have worked (and paid social security contributions) for at least 52 weeks during the two years prior to birth to be eligible to benefits. The 2000 reform introduced the possibility for mothers without prior work history to collect child benefits for up to 30 months. We restrict our sample to women meeting the 52 weeks work requirement to make our estimates comparable across all three reforms.

¹⁹Density tests systematically confirm the absence of manipulation of the timing of births around the three cutoff dates.

²⁰We focus the following analysis on mothers, as fathers' outcomes appear unaffected by PL regimes, as shown in Figure 2. For completeness, we confirm at the end of this section that the estimated effects of the PL reforms for fathers are precisely zero.

4 years before child birth do not feature any discontinuity at the cutoff date . Similarly, panels (b)-(e) show no discontinuity in earnings from event year -3 to event year 0. The absence of any effect in event year 0 (panel (e)) is consistent with the fact that the PL extension from 12 to 24 months did not have any bite in the year of child birth.²¹ The evidence from panels (a)-(e) validate our RD design by showing that there is no selection in earnings around the cutoff date of the reform. Moving to panels (f) and (g), corresponding to event years 1 and 2, we find a strong discontinuity in earnings at the cutoff. Because earnings are measured at the calendar-year level, these two event years correspond to the time during which the parental leave extension had bite.²² The PL extensions cause a significant drop in annual earnings of more than €1,000 in both of these event years. Strikingly, these significant negative effects fully disappear in event year 3, as shown in panel (h). Moreover, we find no sign of any discontinuity in the earnings of mothers in any event year after that, as evidenced by panels (i)-(o).

To confirm the visual diagnostic from Appendix Figure B.I, we estimate the effects of the 1990 reform on earnings Y_t^w for each event year t using the following event-study specification:

$$Y_{it}^w = \alpha D_{it}^{Event} + \alpha^T D_{it}^{Event} \cdot Treat_i + \alpha^B D_{it}^{Event} \cdot Birth_i + \delta D_{it}^{Event} \cdot Treat_i \cdot Birth_i + \beta D_{it}^{Age} + v_{it} \quad (3)$$

To leverage the RD design for identification, we estimate this specification on a sample of mothers who give birth in a narrow window of four months around July 1st, and compare women who gave birth just after the cutoff date ($Treat_i = 1$) to women who gave birth just before the cutoff date ($Treat_i = 0$). We control for seasonality using women who gave birth in a similar four-month window around July 1st in 1989 ($Birth_i = 1989$). The impact of the reform on earnings at each event time t is given by the parameter δ_t . To make magnitudes interpretable, we scale $\hat{\delta}_t$ by the average counterfactual earnings $E[\tilde{Y}_{it}^w | t, Birth_i = 1990]$ obtained from specification (3) when omitting the contribution of the event dummies. The statistic $\hat{\delta}_t / E[\tilde{Y}_{it}^w | t]$ captures the percentage-point change in the child penalty at event time t caused by the reform. The results, reported in Panel (a) of Figure 3, first confirm the absence of selection on pre-child earnings. They also confirm the absence of any significant effect of the reform on

²¹The strong positive slope in the relationship between earnings and the running variable in event year 0 can be understood by noting that earnings are measured at the calendar year level in the ASSD data. Event year 0 therefore corresponds to calendar year 1990. As women who give birth in the earlier months of 1990 spend more time in parental leave in that calendar year, their earnings are mechanically lower.

²²For women who give birth in July 1990, event year 1 corresponds the period going from January to Dec 1991, that is the period going from month 6 to month 18 after birth. Similarly, event year 2 corresponds to the period going from 18 to 30 months after birth.

earnings in event year 0. The negative effect of the 1990 PL extension is entirely concentrated in event years 1 and 2, i.e. the years in which the additional leave entitlement occurs: For both years, the extension from 12 to 24 months of maternal leave causes a 10% drop in female earnings. We find no significant effect of the 1990 reform on female earnings after event year 2. Overall, we find that the 1990 PL extension caused an average 4.04% (0.62%) reduction in the annual earnings of mothers over the first 5 years after the birth of their first child. In Panel (b) of Figure 3, we draw counterfactual child penalties under the pre- and post-1990 PL regimes. We start from the baseline child penalty profile estimated on the sample of mothers who gave birth before the 1990 reform, and then add, for each event year, the estimates of the effects of the reform from panel (a). This gives us the counterfactual child penalty under the 24-month maternal leave regime. The graph shows clearly that extending maternal leave from 12 to 24 months causes a larger child penalty in the short run, but has no impact on the child penalty in the long run.

Using the same empirical strategy, we now turn to the effects of the 1996 reform. This reform reversed part of the 1990 extension by reducing the duration of maternal leave from 24 to 18 months. Visual inspection of the relationship between earnings and the assignment variable in Appendix Figure B.II reveals no sign of a discontinuity at the July 1st 1996 cutoff in any event year, except for event year 2. In this year, we see an increase in earnings of more than €2000 for women who gave birth just after the cutoff date. Panel (c) of Figure 3 shows the associated dynamic RD estimates of the effects of the 1996 reform, which confirm qualitative findings for the 1990 reform. The effects of the reform are fully concentrated in event year 2, i.e. in the year where the parental leave contraction had bite.²³ As before, the reform did not have any long-run effects on the careers of mothers: the increase in labor supply and earnings in event year 2 does not translate into any positive earnings gains in the longer run.

Finally, we consider the 2000 parental leave extension from 18 to 30 months. RD plots for each event year are provided in Appendix Figure B.III and the corresponding dynamic RD estimates are presented in panels (e)-(f) of Figure 3. The analysis confirms the key qualitative insights obtained from the previous reforms: the earnings effect of parental leave is concentrated entirely in the event year directly covered by the PL extension, with no significant effect elsewhere. The only noticeable difference between this reform and the earlier ones is the smaller magnitude of the effect. Compared to the 1990 reform, which also extended the duration of the parental leave by one year (but from a lower baseline), we find that the 2000 reform caused a significantly smaller decline of 2.60%, as compared to 4.04% for the 1990 reform, in the average earnings of mothers over the first 5 years following child birth.

²³For women who gave birth around the July 1st 1996 cutoff, event year 2 is defined as calendar year going from January to December 1998, which corresponds to 18 month to 30 months after birth.

For completeness, Appendix Figure B.IV considers the effect of PL reform on fathers. Since 1990, fathers have been entitled to 6 months of parental leave, but it is possible that the large changes in maternal leave entitlements have had substitution effects on paternal care. Following the same approach as above, we compare fathers of children born on each side of the eligibility cut off for each reform. For all three reforms, we find a precisely estimated zero effect on the trajectory of male earnings.

3.2 The Introduction of Parental Leave in 1961

We now turn to the introduction of paid maternal leave in January of 1961. This reform entitled mothers to receive maternity benefits for up to one year after the birth of a child.²⁴ Maternity benefits were set to the level of unemployment insurance benefits, and receipt was conditional on not working or owning arable land. The maternity leave scheme came with job protection. Empirical analyses of reforms that introduce both job protection and pecuniary benefits have been scant.²⁵ The main reason is that these reforms were implemented a long time ago in most countries, prior to the era usually covered by administrative registry data. In our context, however, we are able to leverage the REV register, which has information on labor market outcomes around the time of the 1961 reform.

To estimate the causal effect of this reform on the earnings dynamics of mothers, we need to account for the fact that the reform was grandfathered. That is, women who gave birth *before* January 1961 were also eligible for the new leave scheme, which they could use up until the first birthday of their child. As a result, scheme eligibility did not display a sharp discontinuity for mothers of children born after January 1961. Indeed, Appendix Figure B.V shows that take-up of maternity benefits featured a gradual increase starting with children born in January 1960. To deal with this issue, instead of the RD approach used above, we analyze the 1961 reform using a difference-in-differences (DiD) approach comparing different birth cohorts over time. Specifically, we compare the outcomes Y_{ist}^w of women whose first child was born between May and August 1961 and who were therefore fully eligible for maternity leave benefits ($T_i = 1$) to women whose first child was born between May and August 1959 and who were fully ineligible ($T_i = 0$). We implement this approach by running the following event-study specification on the sample of women who had their first child in either 1959 or 1961:

$$Y_{it}^w = \alpha D_{it}^{Event} + \alpha^T D_{it}^{Event} \cdot T_i + \beta D_{it}^{Age} + \gamma D_{it}^{Year} + v_{it} \quad (4)$$

²⁴Maternity benefit payments (*Karenzurlaub*) technically start at the end of the maternity insurance period (*Mutterschutzgesetz*), which covers women for 8 weeks following delivery.

²⁵A notable exception is [Schönberg and Ludsteck \(2014\)](#).

Figure 4 shows the time series of DiD estimates $\frac{\hat{\alpha}_t^T}{E[\tilde{Y}_{it}^g|t, T=1]}$, i.e. the percentage-point impact on child penalties for women who became eligible to 12 months of maternal leave. Panel A shows the effects on employment, while panel B shows the effects on earnings. The results indicate that the qualitative effects of the 1961 reform are strikingly similar to those observed for the other parental leave reforms. There is a clear negative effect on both employment and earnings in the first 12 months after birth, i.e. in event years 0 and 1.²⁶ These negative effects quickly disappear as women exhaust their maternal leave entitlement; we observe no significant effects in the long run.

The causal interpretation of these results relies on the assumption that there is no confounding trend in child penalties by birth cohort. To validate this assumption, Appendix Figure B.VI shows placebo results from specification (4) estimated on a sample of mothers who gave birth in 1962 ($T_i = 0$) or 1963 ($T_i = 1$). This figure supports our identifying assumption by showing that there is no significant trend in child penalties across birth cohorts.²⁷

3.3 Why Does Parental Leave Not Hurt Mothers' Careers?

Scholars have expressed mixed views on the long-run effects of parental leave policy on gender inequality. On the one hand, some have voiced concerns that generous parental leave schemes, by keeping mothers out of the labor market for too long, may permanently hurt their careers. This view is predicated on findings that, in the short run, parental leave schemes induce women to stay out of the labor force and reduce labor market earnings (e.g. [Lalive and Zweimüller 2009](#), [Rossin-Slater, Ruhm and Waldfogel 2013](#), [Baum and Ruhm 2016](#)). These short-term work interruptions may translate into longer-term effects through experience channels (such as human capital accumulation and signaling) or through preferences changes. On the other hand, some have argued that job-protected parental leave fosters female labor market attachment, thus leading to positive long-run effects on the career trajectories of mothers. This view has been challenged by findings in [Ejrnæs and Kunze \(2013\)](#), [Schönberg and Ludsteck \(2014\)](#), [Dahl, Løken, Mogstad and Salvanes \(2016\)](#), and [Bailey, Byker, Patel and Ramnath \(2019\)](#), who find either small or negative effects of parental leave schemes in the longer run.

Our findings contribute to this burgeoning literature. An advantage of our study lies in the unique ability to estimate causal effects of parental leave policies over the very long run (about 50 years) and at very different baseline durations (0-3 years). Our analysis

²⁶For women who gave birth between January 1961 and December 1961, the period of the first 12 months after birth spans between the beginning of event year 0 (1961) and the end of event year 1 (1962).

²⁷In a similar vein, we implemented doughnut RD specifications excluding births in 1960, to flexibly account for trends in outcomes by birth cohorts, and found similar results.

show that parental leave has a negative effect on gender gaps in the short run, but a precisely estimated zero effect in the long run. Why do the negative short-run effects not translate into longer-run effects through labor market dynamics? We discuss this question below.

Non-Linear Effects of Parental Leave First, our setting allows us to shed light on non-linearities in the effects of parental leave policies, which are usually held responsible for the differences in estimates across contexts and reforms (see [Rossin-Slater 2017](#)). One key source of non-linearity stems from the interaction between the duration of job protection and the duration of cash benefits ([Lalive, Schlosser, Steinhauer and Zweimüller 2013](#)). As discussed in [Olivetti and Petrongolo \(2017\)](#), it could be that the availability of some job protection, relative to no protection at all, ensures continuity of employment and discourages transitions out of the labor market, while further extensions simply delay return to work without any further gains in employment and earnings.

By combining our estimates of the effects of the 1961, 1990, 1996, and 2000 reforms, we can compute the marginal effect of parental leave duration in the following increments: 2-12 months, 12-18 months, 18-24 months, and 24-30 months. The results are presented in [Figure 5](#), which focuses on female earnings effects during the first five years following child birth. The figure shows that the short-run effect of extending parental leave is highly non-linear: the marginal effect (in absolute value) is monotonically declining in baseline duration. Extending the duration of paid leave from 2 to 12 months has a much larger negative effect than subsequent extensions. Extending leave from 12 to 18 months has a larger negative effect than extending leave from 18 to 24 months, while the extension from 24 to 30 months — beyond the duration of job protection — has virtually no effect even in the short run. The tiny impact of the last extension is driven by low take up. This is consistent with the presence of an interaction effect between job protection and cash benefits: unless it is covered by job protection, the take up of additional paid leave is very small.

Importantly, the non-linearities documented in [Figure 5](#) applies only to the short run, and they are driven entirely by decreasing marginal take-up rates. In the long run, the effects are zero across the entire range of baseline durations. Neither the duration of cash benefits nor the presence of job protection seem to make any difference to the long-run labor market outcomes of women.

Parental Leave & Experience Effects What could explain the absence of longer-term effects of PL policies? In the next section, we investigate the possibility that interactions between institutional child care provision and PL policies drive these results in the Austrian context. But the absence of long-run dynamic effects on mothers' careers still

raises a puzzle: how is it compatible with the presence of significant negative effects in the short run? Experience effects would suggest that spending an additional year out of the labor force would negatively affect future earnings. So are returns to experience in the labor market simply zero for Austrian mothers?

Indeed, a large fraction of mothers work in relatively low-skill occupations in Austria, where experience effects have been estimated to be small. However, Appendix Figure B.VII shows that even among mothers in the top quartile of the pre-birth earnings distribution, PL reforms have no long-term effects. As a result, a more credible explanation is that our estimates identify local average treatment effects among *compliers*. Compliance is likely to be higher among mothers, who do not incur a significant career cost of leave take-up. In other words, when making parental leave decisions, households internalize the potential long-term consequences of these choices for their careers.

The Role of Subsequent Fertility Our estimates of the effects of PL policies do not control for subsequent fertility. In practice, fertility could be a mediator of the long-term effects of parental leave on child penalties, and contribute to mask other dynamic effects of the policy.²⁸ To investigate the mediating role of subsequent fertility, we estimate the dynamic effects of PL reforms on the restricted sample of mothers with only one child – i.e. with a completed fertility of one. Panel A of Appendix Figure B.VIII plots the results for the 1990 reform, and does not reveal any differences in the dynamic patterns compared to the full sample of mothers. This suggests that subsequent fertility does not have any significant mediating impact on the earnings responses to PL extensions.

Previous work by [Lalive and Zweimüller 2009](#) found that the 1990 PL reform had a positive and persistent effect on the likelihood of a second birth. Panel B of Appendix Figure B.VIII replicates this result using as the outcome variable an indicator for having a second child. We estimate that the reform had a persistent effect of 2.5pp on the probability of a second child.²⁹ However, these fertility responses are unlikely to have much of an impact on child penalties. First, the fertility effects are relatively small — the

²⁸Specifically, if longer PL duration were associated with lower (higher) subsequent fertility, this could counteract (amplify) any increase in the long-run penalty from experience effects of longer maternity leave. A small literature has investigated the impact of PL reforms on subsequent fertility. [Dahl, Løken, Mogstad and Salvanes \(2016\)](#) find no effect of longer parental leave on subsequent fertility in the context of Norway. [Malkova \(2018\)](#), on the other hand, finds significant effects on overall fertility in Soviet Russia.

²⁹[Lalive and Zweimüller \(2009\)](#) demonstrate that the Austrian PL reforms not only affected the probability of having a second child, but also the timing of fertility. For both the 1990 and 1996 PL reforms, fertility-timing responses are consistent with incentives created by the parental-leave renewal rule: to qualify for a new parental leave, mothers giving births to a new child — while still on parental leave for their previous child — are *not* subject to the previous-work requirement. The 1990 PL reform created a strong incentive to deliver a second child within the two-year parental leave period of the first child. In contrast, under the one-year parental leave duration of the PL regime prior to the 1990 reform, it was difficult to take advantage of the renewal rule for purely biological reasons.

baseline probability of having a second child is 35% within 3 years and 60% within 10 years. Second, the relevant fertility outcome is completed fertility, including all parities. In Panel C of Appendix Figure B.VIII, we investigate the impact of the reform on the average completed fertility of mothers, measured as of 2017, by weekly bins of the birth date of their first child in an eight-months window around the cutoff date of the 1990 PL reform (Panel C). Interestingly, it turns out that any discontinuity at the 1st of July cutoff disappears, indicating that eligibility to the parental leave extension did not have any significant effect on completed fertility.

It should be noted that women who give birth to their first child on different sides of the reform cutoff date, while facing different PL regimes for their first child, face similar PL regimes (on average) for any subsequent children. As a consequence, our estimates so far have identified only how child penalties respond to the PL regime applicable to the first child. To compute the total effect on child penalties of moving, in steady state, from one PL regime to another, we need to account for the effect of PL extensions for births of higher parity. If this effect is significant, subsequent fertility will mechanically increase the total effect of PL extension on child penalties. To investigate this effect, we consider mothers giving birth to their second, third, and fourth child, replicating our dynamic estimates of the impact of PL reforms on earnings for each parity. We then construct the total steady state impact of a reform by adding the dynamic estimates for all parities, weighted by the average completed fertility and average timing between parities, of women who had their first child around the time of each reform. The results are reported in Appendix Figure B.IX. We find that accounting for the effects of PL reforms for all subsequent births only slightly increases the impact of parental leave in the medium run. We still do not detect any significant effect in the long run. Hence, even when considering their full steady state impact, parental leave reforms have very small effects on maternal careers.

4 The Impact of Child Care Provision on Child Penalties

Over the last sixty years, as parental leave generosity was drastically increased in Austria, so was the provision of public child care. Sixty years ago, nurseries were extremely rare and the majority of municipalities did not have a kindergarten (pre-school). Nowadays, nurseries are common and kindergartens even more so, with more than 90% of all children aged 3-5 enrolled in pre-school care. How did this massive child care expansion affect the labor market outcomes of women relative to men? And did the effects interact with the expansion of parental leave generosity?

4.1 The Expansion of Child Care Provision

Measuring Child Care Provision

We start by constructing long time series of child care supply for each Austrian municipality. The central government collects granular information on all nursery & kindergarten facilities for each municipality and year. The government reports, available since 1988, contain information on the location of each child care institution, its opening hours, its number of certified teachers and their contracts (part-time vs full-time), the legal maximum number of children per certified teacher (which varies by type of institution and is subject to state legislation), and whether the institution is public, private, or operated by the church. The data also includes information on the number of children attending each institution by age and number of hours of attendance. We convert this rich information into two indices that capture the level of child care coverage for children aged 1-2 (nursery care index) and children aged 3-5 (pre-school/kindergarten care index) in each municipality and year:

$$\text{Index 1-2} = 100 \times \frac{\# \text{ FTE Child Care Spots for Children Age 1-2}}{\# \text{ Children of Age 1-2}}$$

$$\text{Index 3-5} = 100 \times \frac{\# \text{ FTE Child Care Spots for Children Age 3-5}}{\# \text{ Children of Age 3-5}}$$

To get the numerator of Index 1-2 (Index 3-5, respectively), we first multiply the number of full-time equivalent teachers in each nursery (kindergarten) by the legal maximum number of children per teacher in the given institution. We weigh these numbers by the opening hours of the institution: a full-time institution gets a weight of 1, while a half-day institution gets a weight of 0.5. We sum the numbers across all nurseries (kindergartens) in each municipality to obtain the number of full-time equivalent child care spots available for children aged 1-2 (3-5). We divide the numerator by the total number of children in the given age group in the municipality. This gives us a measure of the probability that a child of a given age has access to a full-time spot in a child care institution for each municipality and year.³⁰ We also combine the information from our two indices into an aggregate index of child care, labelled Index 1-5:

$$\text{Index 1-5} = 100 \times \frac{\# \text{ FTE Spots for Age 1-2} + \# \text{ FTE Spots for Age 3-5}}{\# \text{ Children of Age 1-5}}$$

Appendix Figure C.I illustrates the massive expansion of child care in Austria since the 1980s. It displays the evolution of the nursery care and kindergarten care indices

³⁰Note that we topcode each index at 100, as some communes offer spots in childcare for more than the number of children in their own municipality.

computed for Austria as a whole. Panel (a) focuses on the nursery care index. It shows that the coverage of nursery care was extremely low in 1988: the probability that a child had access to a full-time spot in a nursery was just 5%. The index increased significantly after that, especially during the 2000s, and is now close to 35%. Panel (b) considers kindergarten care. It shows that the coverage of kindergartens is in general much higher than the coverage of nurseries. Kindergarten coverage was around 55% in 1988, and then increased sharply through the 1990s and early 2000s to reach virtually full coverage.

4.2 The Effect of Local Child Care Expansions on Child Penalties

To assess the impact of child care provision on the dynamics of female labor market outcomes, we start by comparing child penalties for women living in municipalities with high vs low child care coverage. We rank municipalities according to the value of their aggregate child care index (Index 1-5) in 1990, and split the sample by whether municipalities are below or above the median. We then compute, for each subsample, the child penalty for mothers whose first child was born between July 1990 and June 1996, i.e. during the period with 24 months of parental leave. To adjust for differences in the pre-birth characteristics of mothers in different municipalities, we use inverse probability weighting based on pre-birth earnings, pre-birth employment, and age.³¹ The results are presented in Figure 6, which shows that child penalties are significantly lower for mothers living in high-coverage municipalities: their earnings penalty is 5.11pp lower on average during the first five years following child birth. This difference is persistent and stable over the longer run. If we scale the difference in earnings penalties by the difference in child care coverage across the two groups, we obtain a large and significant TOT effect of .158 (.002), suggesting that increasing child care coverage from 0 to 100% would translate into a 16pp reduction in the child penalty.

To interpret this evidence as causal, however, we would need to assume away selection on unobservable characteristics correlated with child penalties. In practice, women may decide to live in municipalities with more child care precisely because they want to invest in their careers after becoming mothers. A closer look at the geographical distribution of child care in 1990 confirms that selection is likely to be an issue. Appendix Figure C.II provides a heat map of the child care index in 1990, showing a stark contrast

³¹Concretely, we predict, using a probit model, the probability π_i of being in a municipality with above median child care provision. We use as predictors the number of days of employment in the year before birth, a set of dummies for the age of the mother at first birth, and the log of individual earnings plus one, $\log(y_{it} + 1)$, for event years $t = -5$ to $t = -1$. Based on these probits, we predict the probability that a municipality has above median child care provision in 1990, and we weight mothers in municipalities with above median coverage by $E(\pi_i) / \hat{\pi}_i$ whereas the weight for mothers in municipalities with below median index is $E(1 - \pi_i) / (1 - \hat{\pi}_i)$.

between the very high coverage rates in and around Vienna and the very low coverage rates in the rest of Austria. The strong urban/rural divide in child care coverage suggests that the cross-sectional comparison of child penalties in Figure 6 is likely to be affected by selection and cannot be interpreted as causal.

Identification Strategy: Local Child Care Expansions

While child care was extremely segmented geographically around 1990, local child care expansions through the 1990s and 2000s allowed the rest of Austria to catch up with Vienna and its surroundings. Importantly, this catch-up process happened at different rates and with different timing across municipalities. Figure 7 maps the change in child care provision for each municipality between 1990 and 2000 in panel (a), and between 2000 and 2010 in panel (b). The figure reveals a tremendous amount of spatial heterogeneity in the intensity and timing of child care expansions. Some municipalities expanded child care in the 1990s, while their neighbors or other geographically similar municipalities expanded child care later, in the 2000s. Some municipalities saw very steep increases, while others saw smaller and more gradual increases.

The variation in local child care expansions offers a promising identification strategy. In principle, we are interested in capturing the macro effects of child care expansions on the dynamics of female labor market outcomes, accounting for potential equilibrium effects, in the spirit of [Duflo \(2001\)](#). To the extent that local labor markets are not fully integrated and that treatment effects do not diffuse much across space, spatial variation in child care expansion across municipalities can meaningfully capture macro effects and is therefore appealing from an identification standpoint.³²

Spatial variation in the expansion of child care may be endogenous to the evolution of female labor market outcomes. For instance, increases in local female employment may push municipalities to expand their supply of institutional child care. To deal with such concerns, we isolate episodes of large and sudden increases in child care provision at the municipal level. Specifically, we focus on events where the nursery care index or the pre-school care index of a municipality increases by 20pp or more in a single year. These events are driven by large supply shocks such as the construction of new facilities, or sizeable expansions of existing facilities and teacher staff. We define a municipality as “treated” if it experiences one such episode. Municipalities that never experienced an increase of at least 20pp in their child care index are assigned to a control group. Our approach is a difference-in-differences (DiD) design in which we

³²Evidence from [Lalive, Landais and Zweimüller \(2015\)](#) indeed suggests that local labor markets are quite geographically segmented in Austria. Studying the response to a large region-specific UI extension program for older workers, they find significant equilibrium effects in treated regions, but very limited geographical spillovers to neighboring municipalities in untreated regions.

compare treated municipalities before and after an episode of child care expansion to control municipalities before and after a placebo episode.³³ To probe whether treated and control municipalities are balanced on observables, Appendix Table C.I provides descriptive statistics by treatment status. The table shows that, prior to local child care expansions, the two groups are remarkably similar in terms of population structure and labor market outcomes.

In Figure 8, we document the evolution of child care provision and child care take-up, i.e. the first stage of our DiD design. Panel (a) plots the evolution of the nursery care index around the time of a nursery care expansion separately for treated and control municipalities. The graph shows that the increase of nursery care provision in treated municipalities has four important characteristics: it is unanticipated, sudden, large, and persistent. By unanticipated, we mean that the index is very stable in the years leading up to the expansion event, and that we do not detect any significant differences in pre-trends between treated and control municipalities. By sudden, large, and persistent, we mean that the index features an immediate jump of 40pp in treated municipalities and remains stable thereafter. For control municipalities, on the other hand, we do not see any noticeable trend in nursery care provision.

As mentioned, one might be concerned that child care expansions are endogenously driven by the dynamics of local demand. An increase in child care demand may create a shortage in supply, prompting local authorities to expand the provision of institutional care. To rule out such concerns, panel (b) of Figure 8 plots the evolution of nursery care take-up around the expansion event. Reassuringly, the graph does not reveal any differential pre-trend in take-up between treated and control municipalities. Furthermore, the figure shows that the dynamics of child care take-up closely mirrors the dynamics of child care provision. The large increase in the provision of nursery care in treated municipalities translates into a large increase in the take-up of nursery care. On average, the supply shocks we study increase the fraction of children going to nursery by a factor of 3.5, from 8% to 28%.

Panels (c) and (d) replicate the analysis for pre-school care. These panels reveal similar dynamic patterns as those observed for nursery care. The main difference relates to the baseline level of child care in treatment vs control municipalities. In the case of nursery care expansions, control municipalities are at a lower baseline than treated municipalities, who are getting further ahead by expanding child care. In the case of pre-school care expansions, treated municipalities are lagging behind and are using the expansions to catch up.

To estimate the effects of child care expansions, we use the following DiD specification.

³³To define these placebo episodes, we randomly assign an expansion event year to control municipalities from the distribution of expansion event years observed in the treatment group.

We compare the evolution of earnings for mothers in treated municipalities ($T_i = 1$) to those in control municipalities ($T_i = 0$) in the years s around a child care expansion, indexing the year of the expansion event to $s = 0$. Because we are interested in mapping out the effects of child care on the full dynamics of labor market outcomes around child birth, our outcomes of interest are the earnings Y_{ist} of mothers observed t years after/until the birth of their first child. For each t , we specify the following model:

$$Y_{ist} = \alpha_t D_{ist}^{Exp} + \alpha_t^T D_{ist}^{Exp} \cdot T_i + \beta D_{ist}^{Age} + \gamma D_{ist}^{Year} + v_{ist} \quad (5)$$

where D^{Exp} is a vector of event time dummies with respect to child care expansion. As before, we control for calendar year fixed effects and age fixed effects. Identification relies on the usual parallel trend assumption that, absent child care expansion, the evolution of maternal earnings t years after birth would have been the same in treated and control municipalities.

For the purposes of estimation, the DiD model (5) for different event times t relative to child birth is stacked into a single specification. That is, we estimate

$$Y_{ist} = \sum_j (\alpha_j D_{isj}^{Exp}) \cdot \mathbf{1}[j = t] + \sum_j (\alpha_j^T D_{isj}^{Exp} \cdot T_i) \cdot \mathbf{1}[j = t] + \beta D_{ist}^{Age} + \gamma D_{ist}^{Year} + v_{ist} \quad (6)$$

The stacked model allows us to calculate counterfactual earnings \tilde{Y}_{ist} , i.e. earnings without the child penalty. We standardize the estimated coefficients $\hat{\alpha}_{st}$ and $\hat{\alpha}_{st}^T$ by $E[\tilde{Y}_{ist} | t]$. To get a visual representation of our DiD estimates, we plot, for each t , the standardized coefficients $\frac{\hat{\alpha}_{st}}{E[\tilde{Y}_{ist}|t]}$ and $\frac{\hat{\alpha}_{st} + \hat{\alpha}_{st}^T}{E[\tilde{Y}_{ist}|t]}$ in a ten-year window around the expansion event ($s = -5$ to $s = 5$). The difference between these coefficients represents the effect on the child penalty t years after birth of being in a treatment municipality (relative to control municipality) as a function of time s relative to a child care expansion. When presenting the results, we set $s = -1$ as our baseline category, so that all effects are expressed relative to the year prior to child care expansion. To account for spatial correlation in unobserved labor market shocks, we cluster standard errors at the municipal level.

The Limited Effects of Child Care Expansions

In Figure 9, we present our results for nursery care expansions. Panel (a) considers women one year before child birth ($t = -1$), and reports earnings impacts in treated municipalities and control municipalities from five years before to five years after the

expansion shock. The graph shows no evidence of differential pre-trends, alleviating any remaining concerns that the supply shock was an endogenous response to the evolution of female labor market outcomes in treated municipalities. The graph also reveals that the expansion had no effect on the earnings of women prior to giving birth. In other words, the increased access to child care did not lead to anticipatory effects among women about to have children.

Panel (b) considers women one to two years after child birth, corresponding to the time where children are enrolled in nursery care. As in panel (a), we do not detect any sign of differential trends before the expansion. Strikingly, we also do not detect any significant difference in the child penalty of women in treated municipalities *after* the expansion. Despite the large increase in the take-up of nursery care, documented in Figure 8, the evolution of earnings one to two years after birth remains virtually identical in treated and control municipalities. The corresponding DiD estimate implies that the nursery care expansions decreased the average child penalty in the first two years following birth by 1.1pp (standard error 1.4pp). If we scale this estimate by the change in nursery care index, we find that increasing municipal nursery care coverage from 0 to 100% would reduce the child penalty by only 3.1pp (standard error 4.0pp).

To corroborate this null result, we replicate DiD graphs similar to those in panels (a) and (b) for child penalties at each event time t , from 3 years before birth to 5 years after birth. The results, reported in Appendix Figure C.III, confirm that the large municipal expansions of nursery care had no detectable effect on child penalties. For every panel, we find that the child penalty estimates in treated and control municipalities are closely aligned over time. In the last panel, we plot the evolution of the average penalty in the 5 years following the birth of a child. These estimates imply that an increase in coverage from 0 to 100% would reduce the child penalty by only 3.0pp (standard error 4.3pp) in the first five years after birth.

A potential concern is that frictions in the labor market, or slow-moving gender norms, significantly hinder the response to nursery care expansion in the short run. However, the results are identical if we look at longer-run effects. In panel (c) of Figure 9, we plot for all event times t the coefficients $\hat{\beta}_{st}$ evaluated at $s = 10$, ten years after the expansion. These estimates capture the long-run causal effect of nursery care expansions on the full dynamics of the child penalty. Note that panel (c) is a mirror image, for nursery care reforms, of panels (a), (c) and (e) in Figure 3 for parental leave reforms. The graph shows without ambiguity that, even in the long run, these large expansions of municipal nursery care supply had no significant effect on child penalties. Based on these estimates, we draw in panel (d) counterfactual child penalties under the pre- and post-regimes: the two profiles are indistinguishable.

One potential explanation for the absence of child care effects is the availability of long

maternal leaves during our period of analysis. Mothers can enrol their child in nursery while still being on maternal leave. The possibility of combining institutional care with maternal leave is no longer available when looking at pre-school years, i.e. when children are between 3 and 5 years old. Yet, Figure 10 shows that episodes of large pre-school expansions had no significant impact on child penalties either. Panels (a) and (b) show that earnings, either just before birth or 3-5 years after birth, evolved in the exact same way in treated and control municipalities. Moreover, looking at the long-run impacts in panels (c) and (d), we find that child penalties were not significantly impacted by large increases in the local supply of pre-school care.

4.3 Interaction Between Child Care and Parental Leave Policies

Given both child care and parental leave policies were found to have a zero effect on child penalties, there should be no significant interaction effect between the two policies either. To confirm this, Appendix Figure C.V shows the effects of the 1990 parental leave extension, comparing women living in municipalities with high and low levels of child care coverage. We first split municipalities by the value of their aggregate index of nursery and pre-school care (Index 1-5) in 1990. Panel (a) shows that the estimated effect of the reform on child penalties is identical for women in municipalities above and below the median of child care provision. Panel (b) focuses specifically on the supply of nursery care, while panel (c) focuses on the supply of pre-school care: in both cases, we do not detect any significant differences in the impact of the parental leave reform between municipalities with high and low provision of child care.

4.4 Why Does Child Care Not Improve Mothers' Careers?

The existing literature on the impact of child care policies on mothers' labor market outcomes has produced mixed results. Some studies estimate positive and sometimes large effects of child care subsidies on female labor supply (e.g., Baker, Gruber and Milligan 2008; Berlinski and Galiani 2007; Cascio 2009; Eckhoff Andresen and Havnes 2018). These micro studies are corroborated by cross-country evidence. At the same time, a number of other studies provide a much less optimistic view on the potential for child care provision to improve female labor supply. Studies from the U.S. (Fitzpatrick 2010, 2012; Barua 2014), Norway (Havnes and Mogstad 2011), France (Goux and Maurin 2010), and Spain (Nollenberger and Rodríguez-Planas 2015) find that child care expansions did not generate any significant improvements in maternal employment. Or if they did, the effects were small and typically limited to single mothers. Our analysis of very large expansions of heavily subsidized child care in Austria shows a

precisely estimated zero effect on mothers' careers, in the short run as well as in the long run.

But why are large increases in the take-up of child care not conducive to increasing female labor supply? Uncovering the mechanisms behind this null effect is critical for understanding the ability of family policies to shape the evolution of gender inequality.

Crowding-Out, Frictions and Preferences Two main mechanisms may explain our findings. First, institutional child care may simply crowd-out informal child care arrangements such as care by relatives. Second, mothers may face large resource and utility costs of increasing labor supply even after child care becomes freely available. These costs may derive from the existence of frictions and constraints in the labor market, or they may be due to preferences and norms regarding maternal care.

To probe the mechanisms behind our results, we use two large surveys conducted with the Austrian Microcensus in 1995 and 2002. These surveys contain household-level information on time use, preferences, and child care arrangements. To investigate the correlation between these outcomes and child care provision, we match the survey information to our measures of institutional child care provision at the political district level.³⁴

To increase statistical power, we pool observations from the two waves of the survey. We control for differences in the education and age composition across districts. Specifically, we regress the average outcome in a political district on the fraction of the population in three education bins and five age bins. We then plot a binscatter of the residuals, rescaled at the average level of education and age in Austria, against the value of the child care index (Index 1-5).

The results are presented in Figure 11. We first consider the relationship between the child care index and the survey measures of child care take-up (panel (a)) and maternal employment (panel (b)). For these two outcomes, the correlation with the child care index can be directly compared to the causal estimates from our local expansion design. The results in panels (a) and (b) shows that the correlations closely match our quasi-experimental estimates. For take-up, we find that a 1 percentage point increase in the child care index translates into a .36 (.09) percentage point increase in the fraction of children attending institutional care. This is remarkably close to the corresponding estimate from our local expansion design: when pooling all nursery and pre-school expansions together, we find a take-up rate of .35. Similarly, both the cross-sectional

³⁴The surveys are unfortunately not representative at the municipal level, but are representative of the 95 Austrian political districts ("*politischer Bezirk*"). These districts correspond to the second most granular level of administrative authority after municipalities.

and local expansion designs are congruent in finding no significant correlation between child care coverage and mothers' probability of employment.

Having established the internal validity of the cross-sectional design, we turn to the correlation between institutional child care provision and households' child care arrangements. Panel (c) considers weekly hours spent on maternal child care. The panel shows that the overall amount of maternal care is very high, with mothers spending about 40 hours per week on average. Moreover, we find no sign that the availability of institutional care reduces the amount of maternal care: mothers spend roughly the same amount of time on child care in districts where the child care index is below 30% and in districts where the index is above 70%. While higher child care provision does not affect maternal care, panel (d) shows that it does crowd-out alternative care by relatives: the fraction of children under the care of relatives decreases by .35 (.11) percentage point when the child care index increases by 1 percentage point. This suggests that working mothers use relatives as child care providers prior to the availability of publicly provided child care, and then switch to nurseries and pre-schools when they become available.³⁵

But what about women who do not work: are they facing constraints that prevent them from increasing labor supply? In panel (e), we correlate child care provision with the fraction of mothers reporting that their career is constrained by a lack of appropriate child care. Only a small fraction of surveyed women say that they feel constrained by the supply of institutional care. What is more, the fraction is no larger in districts with low levels of child care provision than it is in districts with high levels.

If child care constraints are not preventing mothers from improving their career trajectories, then what is? Evidence in panel (f) points to the potentially important role of preferences and norms regarding maternal care: An overwhelming majority of women (70-80%) report that they do not work, because they have a preference for taking care of their children. Furthermore, we find only a marginally significant correlation between these reported preferences and the level of child care provision. This indicates that expanding access to institutional child care has not shifted the strong norms regarding maternal care.

Overall, these results suggest that the combination of cheap available care by relatives and strong preferences for maternal care are the main reasons why the large expansion of child care provision in Austria has had no effect on female labor market outcomes.

³⁵In a recent study for Austria, [Frimmel, Halla, Schmidpeter and Winter-Ebmer \(2020\)](#) find that grandmothers' labor supply reacts to the presence of grandchildren. A timing-of-events approach shows that a first grandchild increases the probability of leaving the labor market by 9 percent, an effect which is particularly strong when grandmothers live close to the grandchild.

Fertility Effects To fully understand the impact of child care expansions on female careers, we finally need to account for the possibility that the expansions affected fertility. If the increased availability of child care induced women to have more children, this may have mechanically increased child penalties, all else being equal. That is, the null effect of child care expansions could reflect offsetting effects of smaller child penalties at a given level of fertility and larger child penalties due to fertility responses. The raw time series correlations suggest against such an interpretation: The roll-out of cheap institutional child care in Austria over the last sixty years has been accompanied by a steady decline in fertility rates. Of course, these time series patterns may be affected by confounders that mask any causal effect of child care provision on fertility.

The existing literature offers relatively little guidance on this question. Few studies provide clear evidence on the fertility effects of child care provision.³⁶ Fortunately, our setting allows for estimating the effect of municipal child care expansions on local fertility using a DiD approach. To this end, we measure local fertility rates as the total number of births divided by the total number of women aged 15-49 in a municipality. Appendix Figure C.VI plots the evolution of fertility rates in treated and control municipalities around the time of large shocks in child care supply. Looking at the long-run effects, ten years after the expansion, we find small and statistically insignificant effects of child care expansions on fertility. This is true for both nursery and pre-school care. These findings suggest that our null effect of child care expansions are not the result of offsetting fertility effects.

5 How Much of the Evolution of Gender Inequality Can Family Policies Explain?

Equipped with causal estimates of the effects of all family policy reforms in Austria over the last sixty years, we can now measure how the dramatic expansion of family policies has shaped the evolution of gender inequality. To do this, we proceed in two steps. First, we use the estimated child penalties to decompose gender inequality into child-related and residual inequality. We can then, in a second step, leverage our estimates of the effects of family policies on child penalties to simulate gender inequality under alternative policy scenarios. This will uncover how family policies have affected the long-run historical evolution of gender inequality.

³⁶An exception is [Bauernschuster, Hener and Rainer \(2016\)](#), who find significant fertility effects of local child care expansions in Germany.

5.1 Child-Related Gender Inequality Over the Last 60 Years

The consistency and quality of the Austrian administrative data allows us to reconstruct the evolution of gender gaps in earnings since the beginning of the 1950s. For this purpose, we combine the digitized ASSD data starting in 1972 with the detailed tabulated statistics on the distribution of earnings by gender based on the same underlying ASSD data before it was computerized (*Lohnstufenstatistik*).³⁷ This allows us to measure, for all years s between 1953 and 2017, the total gender gap in raw earnings $\Delta_s \equiv 1 - E[Y_{ist}^w|s]/E[Y_{ist}^m|s]$. Two points should be stressed about our measure of the gender gap. First, it corrects for top coding in the data by interpolating the top tail of the distribution using information from uncensored tax data.³⁸ Second, we account for the extensive margin of labor supply by including individuals with zero earnings.³⁹ In that sense, our measure of the gender gap encompasses the broadest possible set of behaviors underlying labor market outcomes.

For each year, we can calculate how much of the cross-sectional gender gap Δ_s is explained by child penalties, i.e., by the differential impact of children on the careers of mothers and fathers. To do this, we follow a decomposition methodology similar to the one developed by [Kleven, Landais and Sogaard \(2019\)](#). This strategy relies on using child penalty estimates to get a measure of the counterfactual gender gap $\tilde{\Delta}_s \equiv 1 - E[\tilde{Y}_{ist}^w|s]/E[Y_{ist}^m|s]$. This counterfactual captures the level of gender inequality in earnings that would prevail if we removed the negative effect of children on the careers of women relative to men. Child-related gender inequality is computed as the difference between the observed and counterfactual gender gaps:

$$\Delta_s - \tilde{\Delta}_s \equiv \frac{1}{E[Y_{ist}^m|s]} \cdot \left[\sum_t \psi_{st} \cdot P_{st} \cdot E[\tilde{Y}_{ist}^w|s, t] \right] \quad (7)$$

In the language of Oaxaca-Blinder decompositions, our measure of child-related inequality corresponds to the “unexplained” effects of children and two sets of inputs are necessary to compute it. First, we need a measure of the fraction of women ψ_{st} who are at a given event time t relative to the birth of their first child in calendar year s . Second, we need estimates of their child penalty P_{st} and of their corresponding average counterfactual earnings $E[\tilde{Y}_{ist}^w|s, t]$.

We compute ψ_{st} by complementing the ASSD data with information from the birth register and from the Austrian Birth Barometer in order to account for women whose

³⁷See [Christl and Wagner \(1981\)](#). We provide all the details of the data sources and construction of our gender gap measures in Appendix [D.1](#).

³⁸In practice, we allow for different Pareto parameters by gender and by age group. Details of our procedure are given in Appendix [D.1](#).

³⁹As explained in Appendix [D.1](#), we made sure that the fractions of men and of women with zero earnings in our data matched the Census data.

first birth is not observed in the ASSD or REV data. To get estimates of child penalties and counterfactual earnings by year, we extend the baseline specification (1) to allow for year-specific coefficients on event time. That is, for women we estimate the following model:

$$Y_{ist}^w = \sum_y (\alpha_y^w D_{ist}^{Event}) \cdot \mathbf{I}[y = s] + \beta^w D_{ist}^{Age} + \gamma^w D_{ist}^{Year} + \nu_{ist}^w \quad (8)$$

We interact the event time dummies with year dummies to estimate year-specific event coefficients $\hat{\alpha}_{st}^w$. Counterfactual earnings for mothers at event time t and year s simply correspond to the predicted earnings from the above specification when omitting the contribution of the coefficients $\hat{\alpha}_{st}^w$.⁴⁰

Panel (a) of Figure 12 presents the results from our decomposition. The first striking observation is that, despite a significant decline in gender inequality of about 30pp over the last sixty years, the gender gap in earnings remains very large in Austria: the average woman earns 40% less than the average man in 2017. What is more, the level of child-related gender inequality is also very large: women earn on average 30-35% less than men *because of children*. This reflects the large child penalties experienced by Austria mothers. Interestingly, child-related gender inequality has been quite stable over the past six decades. As a consequence, parenthood went from being responsible for just about half of gender inequality at the beginning of the 1950s to being responsible for almost the all of gender inequality today. This implies that most of the historical gender convergence in the labor market has been due to factors that are not directly related to children such as the closing of the educational gender gap.

While the overall level of gender inequality is considerably higher in Austria than in Scandinavia, we note that results from Figure 12 are strikingly similar to the decomposition evidence from Kleven, Landais and Sogaard (2019) for Denmark. They also found that child penalties are the main driving force behind gender inequality today. We now turn to investigating how family policies have influenced the long-run historical evolution of gender inequality.

5.2 Gender Gaps Under Counterfactual Family Policies

Our focus is on estimating how gender inequality has been affected by the dramatic expansion of three family policies: parental leave (τ_{PL}), free nursery care (τ_N), and free

⁴⁰Note that estimating event coefficients by calendar year s and event year t amounts to estimating event coefficients by birth cohort $c = s - t$. For births cohorts before the start of our data, we observe only post-event years and are unable to directly estimate child penalties associated with these births. We therefore rely on an extrapolation of the observed penalties in the data that we describe in detail in Appendix D. Overall, results reported in Appendix Figure D.VII show that penalties up to 10 years after birth have been remarkably stable over the past 50 years. Longer run penalties, from 10 to 30 years after birth, to the contrary have declined significantly.

pre-school care (τ_{PS}). This means identifying, for each policy $\tau_k \in \{\tau_{PL}, \tau_N, \tau_{PS}\}$, the causal effect of the policy on the gender gap, $\frac{d\Delta_s}{d\tau_k}$. We can express this as a function of the causal effect on the dynamics of child penalties, $\frac{dP_t}{d\tau_k}$ at each t . That is, we have

$$\begin{aligned} \frac{d\Delta_s}{d\tau_k} &= -\frac{dE[Y_{ist}^w|s]/d\tau_k}{E[Y_{ist}^m|s]} \\ &= -\frac{1}{E[Y_{ist}^m|s]} \cdot \left\{ \sum_t \psi_{st} \cdot \frac{dP_t}{d\tau_k}(\tau_{PL}, \tau_N, \tau_{PS}) \cdot E[\tilde{Y}_{ist}^w|s, t] \right\} \end{aligned} \quad (9)$$

where we use the empirical fact that the earnings of men are unaffected both by the arrival of kids and by family policies. Motivated by our empirical evidence, the above expression assumes that family policies have no effect on fertility such that $d\psi_{st}/d\tau_k = 0$. But we explore below scenarii in which we relax this assumption and allow family policies to affect fertility as well. We also note that the effect of any given policy change $d\tau_k$ on child penalties is in general a function of *all* family policies ($\tau_{PL}, \tau_N, \tau_{PS}$). However, we can leverage our empirical finding of no interaction effects between parental leave and child care provision to assume away such complications. Hence, we consider the impact of a given policy change on the child penalty to be a function only of that particular policy, i.e. $\frac{dP_t}{d\tau_k}(\tau_k)$. This, in turn, allows us to easily explore counterfactual scenarii where we change all three policies together.

The scenarii we study involve large, non-marginal policy changes. Our unique ability to measure their impact on the evolution of aggregate gender inequality relies on two fundamental assets. The first is the fact that we have estimated the effect of each policy on the full dynamics of child penalties, i.e., $\frac{dP_t}{d\tau_k}(\tau_k)$ for all event times t . The second is that we can account for non-linearity in the effects of these policies. For parental leave policies, we have identified the effect of changing the duration of PL benefits at different baseline durations, starting from no parental leave at all. For child care provision, our estimates show that even massive reforms — increases in the level of coverage of more than 40pp — do not have any significant effects on child penalties in both the short and long run. This suggests that non-linearities are not important for the provision of child care.

Gender Inequality Without Parental Leave and Institutional Child Care Our baseline counterfactual policy scenario consists in getting rid of all family policies. In other words, we simulate the evolution of the gender gap since 1961 when eliminating the effects of all parental leave reforms and setting the level of nursery and pre-school care provision to zero. The results are presented panel (b) of Figure 12. The graph conveys a clear and striking message: In a world without the massive expansion of family poli-

cies over the last sixty years, gender inequality would have been just about the same. If anything, inequality would have been smaller than it is today. In 2010, the gender gap in earnings would have been 41.4% absent family policies compared to the observed gap of 43.7%. The gender gap caused by children would have been 30.1% instead of the observed gap of 34.7%.

To provide additional insight, Appendix Figure D.VIII shows computations of counterfactual gender gaps when removing each family policy separately. Panel (a) shows the evolution of the gender gap absent any parental leave expansions, assuming that child care provision follows the same path as actually observed. The graph highlights that parental leave expansions on their own have increased gender inequality, although the effect is small. In the absence of parental leave, we find that gender inequality in earnings would have been about 3pp smaller today. This result stems from the negative impact of parental leave policies on earnings in the short run. The absence of long-run effects on the careers on women renders the total impact of parental leave policies on gender inequality modest. While parental leave policies do have some effect on the evolution of the gender gap, panel (b) and (c) shows that child care policies have had absolutely no effect.

These results reflect the small point estimates of the impact of family policies on child penalties. But how sensitive are our counterfactual gender gaps to the precision with which these null effects are estimated? Can we reject that family policies have had large effects on gender inequality? To provide some guidance for inference, we compute upper and lower bounds on counterfactual gender gaps. The first “optimistic” scenario uses, instead of our point estimates, the upper bound of the 95% confidence interval on $\frac{dP_t}{d\tau_k}(\tau_k)$ for each policy change. The second “pessimistic” scenario uses the lower bound of the 95% confidence interval. The results of these two bounding scenarios are presented in Table 1. The bottom line is that, even in the optimistic scenario, family policies do not cause any substantial reduction in the gender gap. And the difference between the optimistic and pessimistic scenarios is minor (.428 vs .400). Our null effects are therefore precise: we can rule out large positive effects of family policies on gender inequality.

Robustness to Equilibrium Effects Our mapping between family policies and aggregate gender inequality rests on our quasi-experimental estimates of the effects of these policies on child penalties. It is therefore legitimate to ask if these estimates fully account for the mechanisms through which these policies could affect gender gaps.

One potential concern relates to equilibrium effects in the labor market: do these estimates properly account for all equilibrium effects that may arise from large family policy changes? For child care provision, our quasi-experimental strategy was aimed precisely at capturing equilibrium effects by focusing on differential treatment *across*

local labor markets. The fact that we found zero impacts on labor supply and earnings implies no equilibrium impact at the level of local labor markets. For parental leave policy, our quasi-experimental strategy compared treated and untreated mothers within the same labor market, and therefore did not address potential general equilibrium effects. Still, the fact that we found very limited partial-equilibrium effects of these policies on child penalties greatly limits the scope for any significant general-equilibrium effects.

Another concern may be the presence of long-run effects operating through slow-moving changes in preferences or norms. Can our estimates capture them adequately? We note that our difference-in-differences evidence reveals no effects for up to ten years after policy implementation. Moreover, our investigation in Figure 11 suggests that preferences for maternal care remain extremely strong and do not react much to the level of child care provision. It is therefore unlikely that our counterfactual exercise misses significant effects of family policies mediated by changes in preferences.

Anticipatory Effects and Fertility Responses We have shown that family policies have been largely irrelevant for the evolution of gender inequality over the last sixty years, because their impact on child penalties are extremely muted. As a consequence, the historical decline in gender inequality must have been driven by factors unrelated to children, or to behavioral changes made before the arrival of kids, such as educational investments and discrimination. This, in turn, opens two final questions.

First, could family policies have affected the choices made by women before the arrival of children, thus contributing to the reduction in gender inequality? For example, when child care became more widely available, did women respond by investing significantly more in their career prior to having children? Could our decomposition therefore miss a key mechanism through which family policies reduce the gender gap? The likely answer is no. Our analysis of local child care expansion in Figures 9 and 10 shows that large increases in child care provision have no anticipatory effect on the labor market outcomes of women prior to having children. Even ten years after local child care expansions, we found no significant differences in employment or earnings for women who were about to have kids in treated and control municipalities.

Even if family policies do not generate any anticipatory effects prior to motherhood, they could still affect the probability of motherhood as analyzed in sections 3.3 and 4.4. This raises our second question: how robust are our counterfactual estimates to accounting for fertility responses? What has been the causal effect of the expansion of family policies on fertility? And what were the consequences for gender inequality?

As discussed above, we have very limited evidence on the impact of family policies on fertility and thus no consensus. But as far as child care provision is concerned, we

can use the results from our local child care expansions in Appendix Figure C.VI. There we found small and insignificant fertility responses. For parental leave, the evidence is even more scant. Most studies in the literature have compared, like we did in Appendix Figure B.VIII, the completed fertility rates of women who are exposed, for their first child, to short vs long parental leave durations. But this does not capture the relevant effect on total fertility, because in such a quasi-experimental setup, women face the same parental leave regime for any future children. Furthermore, one cannot identify from such variation the effect of parental leave on the probability of having a first child. To get at the effect of interest, we would need to compare the total fertility of women who are experiencing two different parental regimes *permanently*. This is challenging because parental leave policies usually apply to all women in a given country. This has forced researchers to rely on cross-country variation as in Olivetti and Petrongolo (2017). Correlating parental leave generosity with fertility rates for 22 countries over the period 1970–2010, controlling for country and year fixed effects, they find a significant, but small, positive correlation.

Using the estimates from Olivetti and Petrongolo (2017) for parental leave combined with our quasi-experimental estimates for child care provision, we can simulate counterfactual gender gaps when allowing for fertility responses.⁴¹ The results are reported in Appendix Figure D.X. This figure shows that fertility responses to family policies are too small to make any substantial difference to gender inequality. We estimate that, when accounting for fertility responses, the gender gap would be 0.423 today in the absence of any family policy. This is almost undistinguishable from our baseline estimate of 0.414 without fertility responses.⁴²

6 Conclusion

In this paper, we have measured the contribution of family policies to the long-run evolution of gender inequality, leveraging historical administrative data and quasi-experimental estimates over more than sixty years. Our approach builds on causally identified estimates of the impact of policy reforms, making it more credible than cross-country approaches to assess the importance of family policies for gender inequality (Ruhm 1998; Blau and Kahn 2013; Olivetti and Petrongolo 2017). By accounting for the

⁴¹We take estimates from Table 3, column (8) of Olivetti and Petrongolo (2017). Appendix D.3 provides all details on our methodology.

⁴²In Appendix Table D.III, we also explore the sensitivity of results to the precision of our estimates of fertility response. We use a methodology similar to Table 1, and compute the gender gap in two scenarios: one using only the upper bound of the 95% confidence interval of our estimates of fertility responses, and the other using only the lower bound of the confidence interval. We find that the range of possible gender gap estimates between these two bounds is very narrow, going from 0.437 to 0.408. We can therefore credibly rule out the presence of significant effects of family policies on gender inequality.

effects of all family policy variation, our approach is robust to policy interactions and other non-linearities that normally provides threats to out-of-sample extrapolations of marginal treatment effects. This general approach could be fruitfully extended to other contexts such as, for example, measuring the contribution of tax and transfer policies to income and wealth inequality.

Our results point to a limited role for public policy in explaining the secular decline in gender inequality. Our setting focuses on one particular country, Austria, but their historical expansion of family subsidies is similar to the policy path of other European countries, and while the U.S. has not gone as far as Europe, proposals to expand family policies are the subject of ongoing debate. We find that parental leave policies have had negative short-term effects on female labor market outcomes, but no long-term effects. The short-term effects are relatively small and largely inconsequential for the evolution of aggregate gender inequality. More surprisingly, we also find that publicly provided and heavily subsidized child care has had no effect on the secular decline in gender inequality. This stands in contrast to the strong cross-country correlation between child care subsidization and gender gaps, while within-country quasi-experimental studies of specific reforms have led to mixed results.

Broadly speaking, our evidence is consistent with the ideas exposed in [Goldin \(2014\)](#), arguing that gender convergence relies less on government intervention than on equilibrium features of the labor market. Our work is also consistent with the notion that gender convergence reflects the evolution of gendered preferences and norms ([Bertrand 2011](#); [Kleven, Landais, Posch, Steinhauer and Zweimuller 2019](#); [Boelmann, Raute and Schönberg 2020](#)). This underlines the need for furthering our understanding of the formation of preferences and social norms regarding the family-career choices of men and women, and to account for their potential interaction with public policies. Endogenous preferences may create multiple equilibria and tipping points, whereby small changes in policies can have large effects on gender inequality. Conversely, as seems to be the case in Austria, gendered preferences can be sufficiently strong that gender gaps do not respond to even very large expansions of policies.

References

- BAIERL, A. AND M. KAINDL (2011): “Kinderbetreuung in Österreich,” Working Paper 77, Austrian Institute for Family Studies, Vienna.
- BAILEY, M. J., T. S. BYKER, E. PATEL AND S. RAMNATH (2019): “The Long-Term Effects of California’s 2004 Paid Family Leave Act on Women’s Careers: Evidence from U.S. Tax Data,” Working Paper 26416, National Bureau of Economic Research, Series: Working Paper Series.

- BAKER, M., J. GRUBER AND K. MILLIGAN (2008): "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 116(4), 709–745.
- BARUA, R. (2014): "Intertemporal Substitution in Maternal Labor Supply: Evidence Using State School Entrance Age Laws," *Labour Economics*, 31, 129–140.
- BAUERNSCHUSTER, S., T. HENER AND H. RAINER (2016): "Children of a (Policy) Revolution: The Introduction of Universal Child Care and its Effect on Fertility," *Journal of the European Economic Association*, 14(4), 975–1005.
- BAUM, C. L. AND C. J. RUHM (2016): "The Effects of Paid Family Leave in California on Labor Market Outcomes," *Journal of Policy Analysis and Management*, 35(2), 333–356.
- BERLINSKI, S. AND S. GALIANI (2007): "The Effect of a Large Expansion of Pre-Primary School Facilities on Preschool Attendance and Maternal Employment," *Labour Economics*, 14(3), 665–680.
- BERTRAND, M. (2011): "New Perspectives on Gender," in *Handbook of Labor Economics*, vol.4b, ed. by O. Ashenfelter and D. Card, chap. 17, pp. 1543–1590. North Holland: Elsevier Science Publishers.
- BERTRAND, M., C. GOLDIN AND L. F. KATZ (2010): "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2, 228–255.
- BLAU, F. AND L. KAHN (2013): "Female Labor Supply: Why Is the United States Falling Behind?," *American Economic Review: Papers and Proceedings*, 103(3), 251–256.
- BOELMANN, B., A. RAUTE AND U. SCHÖNBERG (2020): "Wind of Change? Cultural Determinants of Maternal Labor Supply," Working Papers 914, Queen Mary University of London, School of Economics and Finance.
- CASCIO, E. U. (2009): "Maternal Labour Supply and the Introduction of Kindergartens into American Public Schools," *The Journal of Human Resources*, 44(1), 140–170.
- CHRISTL, J. AND M. WAGNER (1981): "Geschlechtsspezifische Einkommensunterschiede: Österreich 1953-1979," Forschungsbericht, Institut für höhere Studien (IHS), Vienna.
- CORNELISSEN, T., C. DUSTMANN, A. RAUTE AND U. SCHÖNBERG (2018): "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance," *Journal of Political Economy*, 126(6), 2356–2409.
- CORTÉS, P. AND J. PAN (2020): "Children and the Remaining Gender Gaps in the Labor Market," Working Paper 27980, National Bureau of Economic Research.

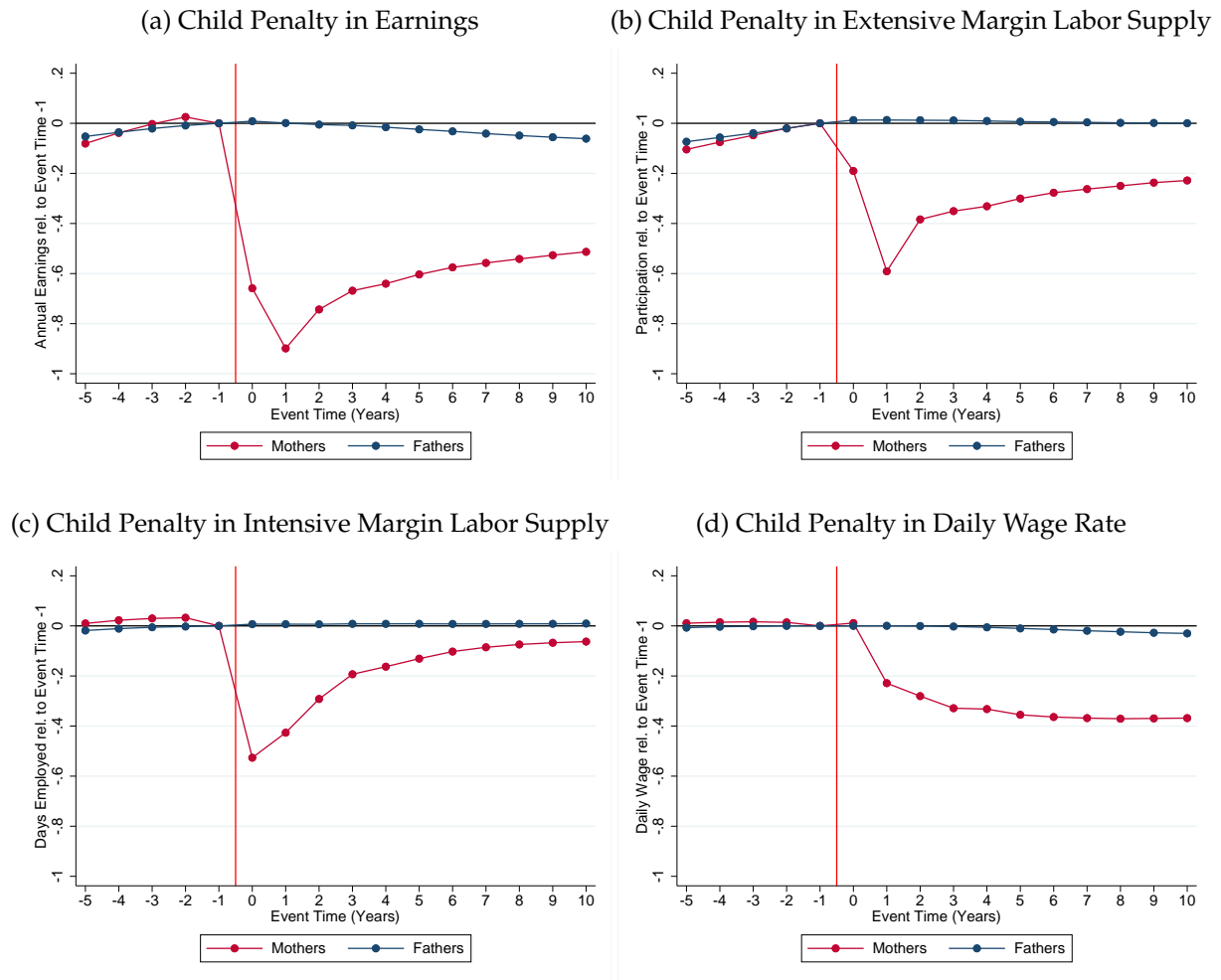
- DAHL, G. B., K. V. LØKEN, M. MOGSTAD AND K. V. SALVANES (2016): "What Is the Case for Paid Maternity Leave?," *Review of Economics and Statistics*, 98(4), 655–670.
- DUFLO, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), 795–813.
- ECKHOFF ANDRESEN, M. AND T. HAVNES (2018): "Child Care, Parental Labor Supply and Tax Revenue," Discussion paper, Institute for the Study of Labor (IZA).
- EJRNÆS, M. AND A. KUNZE (2013): "Work and Wage Dynamics around Childbirth," *The Scandinavian Journal of Economics*, 115(3), 856–877.
- FITZPATRICK, M. D. (2010): "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten," *Journal of Labor Economics*, 28(1), 51–85.
- (2012): "Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool," *Journal of Human Resources*, 47(3), 583–612, Publisher: University of Wisconsin Press.
- FRIMMEL, W., M. HALLA, B. SCHMIDPETER AND R. WINTER-EBMER (2020): "Grandmothers' Labor Supply," *Journal of Human Resources*, p. forthcoming, Publisher: University of Wisconsin Press.
- GOLDIN, C. (2014): "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 104, 1091–1119.
- GOLDIN, C. AND L. F. KATZ (2016): "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation," *Journal of Labor Economics*, 34, 705–746.
- GOUX, D. AND E. MAURIN (2010): "Public School Availability for Two-Year Olds and Mothers' Labour Supply," *Labour Economics*, 17(6), 951–962.
- HAVNES, T. AND M. MOGSTAD (2011): "Money for Nothing? Universal Child Care and Maternal Employment," *Journal of Public Economics*, 95(11-12), 1455–1465.
- KLEVEN, H. (2014): "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 28, 77–98.
- KLEVEN, H. AND C. LANDAIS (2017): "Gender Inequality and Economic Development: Fertility, Education, and Norms," *Economica*, 84, 180–209.
- KLEVEN, H., C. LANDAIS, J. POSCH, A. STEINHAEUER AND J. ZWEIMULLER (2019): "Child Penalties Across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 109, 122–126.

- KLEVEN, H., C. LANDAIS AND J. E. SØGAARD (2019): "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 11(4), 181–209.
- (2020): "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families," *American Economic Review: Insights*, p. forthcoming.
- KUZIEMKO, I., J. PAN, J. SHEN AND E. WASHINGTON (2018): "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?," NBER Working Paper No. 24740.
- LALIVE, R., C. LANDAIS AND J. ZWEIMÜLLER (2015): "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 105(12), 3564–3596.
- LALIVE, R., A. SCHLOSSER, A. STEINHAEUER AND J. ZWEIMÜLLER (2013): "Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits," *Review of Economic Studies*, 81(1), 219–265.
- LALIVE, R. AND J. ZWEIMÜLLER (2009): "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments," *The Quarterly Journal of Economics*, 124(3), 1363–1402.
- MALKOVA, O. (2018): "Can Maternity Benefits Have Long-Term Effects on Childbearing? Evidence from Soviet Russia," *The Review of Economics and Statistics*, 100(4), 691–703.
- NOLLENBERGER, N. AND N. RODRÍGUEZ-PLANAS (2015): "Full-Time Universal Child-care in a Context of Low Maternal Employment: Quasi-Experimental Evidence from Spain," *Labour Economics*, 36, 124–136.
- OLIVETTI, C. AND B. PETRONGOLO (2016): "The Evolution of Gender Gaps in Industrialized Countries," *Annual Review of Economics*, 8, 405–434.
- (2017): "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries," *Journal of Economic Perspectives*, 31, 205–230.
- ROSSIN-SLATER, M. (2017): "Maternity and Family Leave Policy," Discussion paper, National Bureau of Economic Research.
- ROSSIN-SLATER, M., C. J. RUHM AND J. WALDFOGEL (2013): "The Effects of California's Paid Family Leave Program on Mothers? Leave-Taking and Subsequent Labor Market Outcomes," *Journal of Policy Analysis and Management*, 32(2), 224–245.

- RUHM, C. J. (1998): "The Economic Consequences of Parental Leave Mandates: Lessons from Europe," *Quarterly Journal of Economics*, 113(1), 285–317.
- SCHÖNBERG, U. AND J. LUDSTECK (2014): "Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth," *Journal of Labor Economics*, 32(3), 469–505.
- ZWEIMÜLLER, J., R. WINTER-EBMER, R. LALIVE, A. KUHN, J.-P. WUELLRICH, O. RUF AND S. BÜCHI (2009): "Austrian Social Security Database," Working Paper 410, Institute for Empirical Research in Economics (IEW), Zurich.

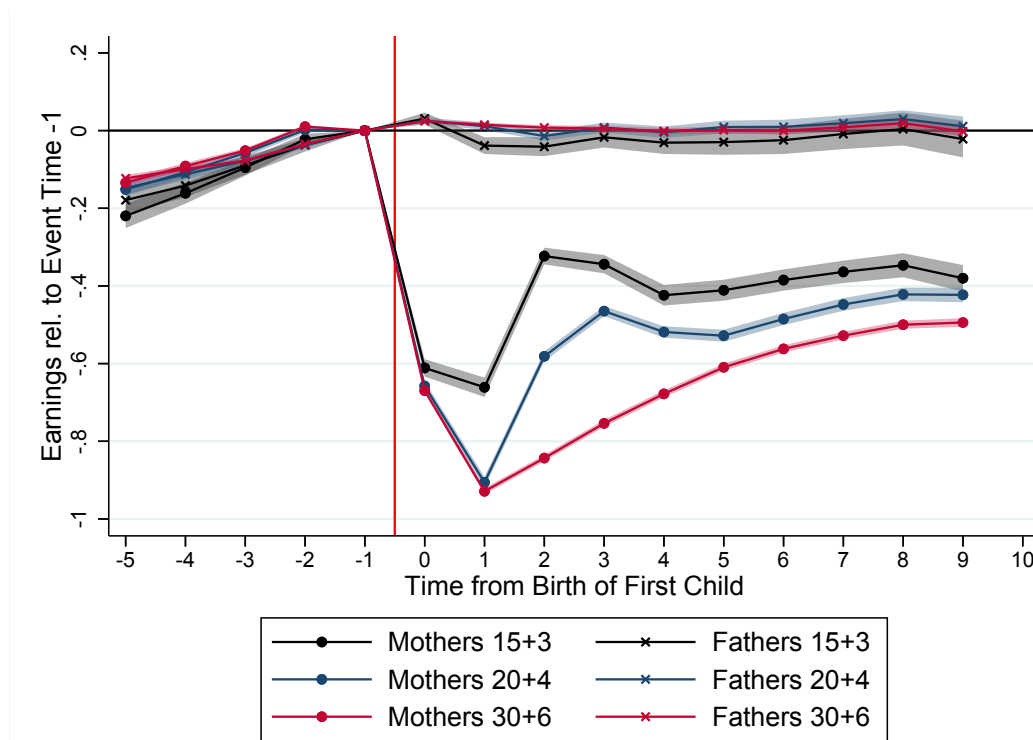
Figures

Figure 1: Child Penalties



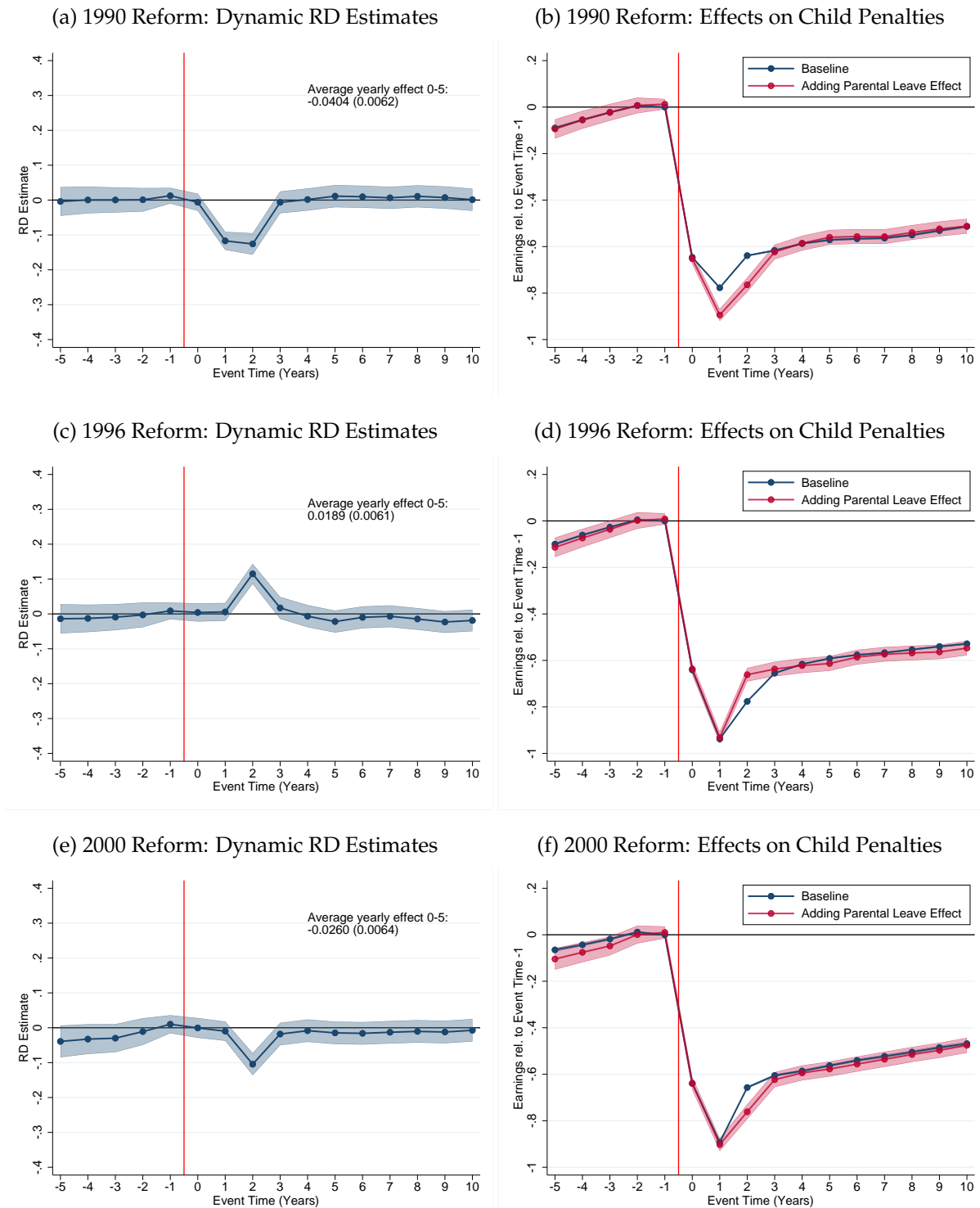
Notes: The figure shows event time coefficients estimated from equation 2 and normalized as explained in Section 2.3 (i.e. we plot $P_t^g \equiv \hat{\alpha}_t^g / E[\tilde{Y}_{ist}^g | t]$) for mothers and fathers and for different outcomes. The sample includes all first births between 1985-2012. Fathers are matched to mothers based on applications for child benefits. In panel (a), the dependent variable is annual earnings (including individuals with zero earnings) converted to real 2000 Euro using the CPI and adjusted for topcoding (see Appendix D.1 for details). In panel (b) the dependent variable is an indicator for whether an individual had any employment spells in a given year. In panel (c), we focus on total days of employment conditional on having at least one employment spell that year. In panel (d), daily wage rate is defined as annual earnings divided by days of employment in that year, and is conditional conditional on having at least one employment spell that year.

Figure 2: The 2008 Parental Leave Regime: Child Penalty by PL Option



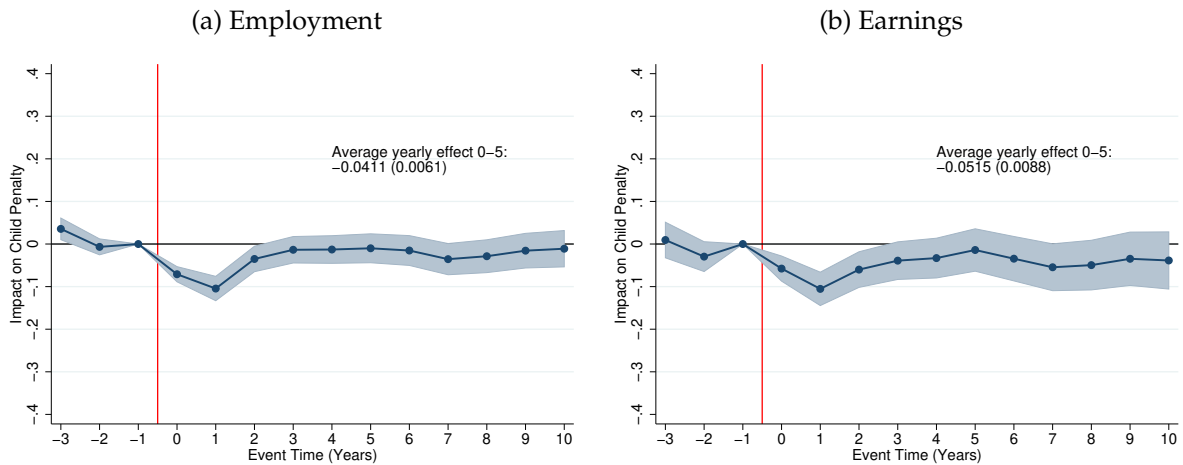
Notes: This figure shows normalized event time coefficients ($P_t^g \equiv \hat{a}_t^g / E[\tilde{Y}_{ist}^g | t]$) from child penalty regressions in annual earnings. The sample is mothers and fathers who had their first child in 2008. We separately estimate child penalty regressions for mothers and fathers according to the parental leave option they selected into. The “30+6” option corresponds to the long leave, with up to 30 months for the mother, and 6 months for the father, at low replacement rate. The “20+4” corresponds to the intermediate option of 20 months of PL for the mother and 4 months for the father. The “15+3” option, corresponds to the short parental leave at higher replacement rate. Note that we assign parents to the different options based on the observed length of their effective leave. Shaded areas are 95% confidence intervals on the normalized event time coefficients.

Figure 3: Parental Leave Reforms: Dynamic RD Estimates & Causal Effects on Child Penalties



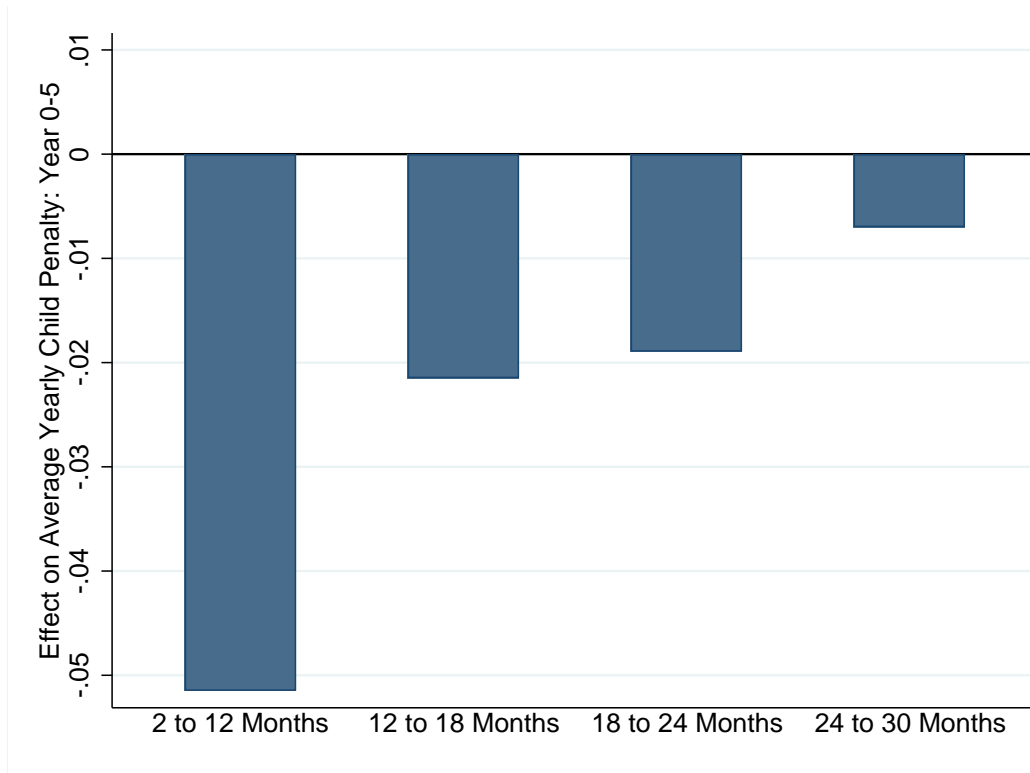
Notes: The figure reports the estimated effects of the 1990, 1996 and 2000 reforms on child penalties in earnings for women. The left panels plots for each event time the statistic $\hat{\delta}_t/E[\tilde{Y}_{it}^w | t]$ estimated from specification 3, which captures the percentage-point change in the child penalty at event time t caused by the reform. The right panels plot the observed child penalties for women whose first child is born before the reform (baseline), and the counterfactual child penalties that these mothers would have had under the new PL regime, which corresponds to adding the estimates from the corresponding left panel to the baseline penalty. Shaded areas correspond to 95% confidence intervals.

Figure 4: Dynamic Effects of the 1961 Parental Leave Reform



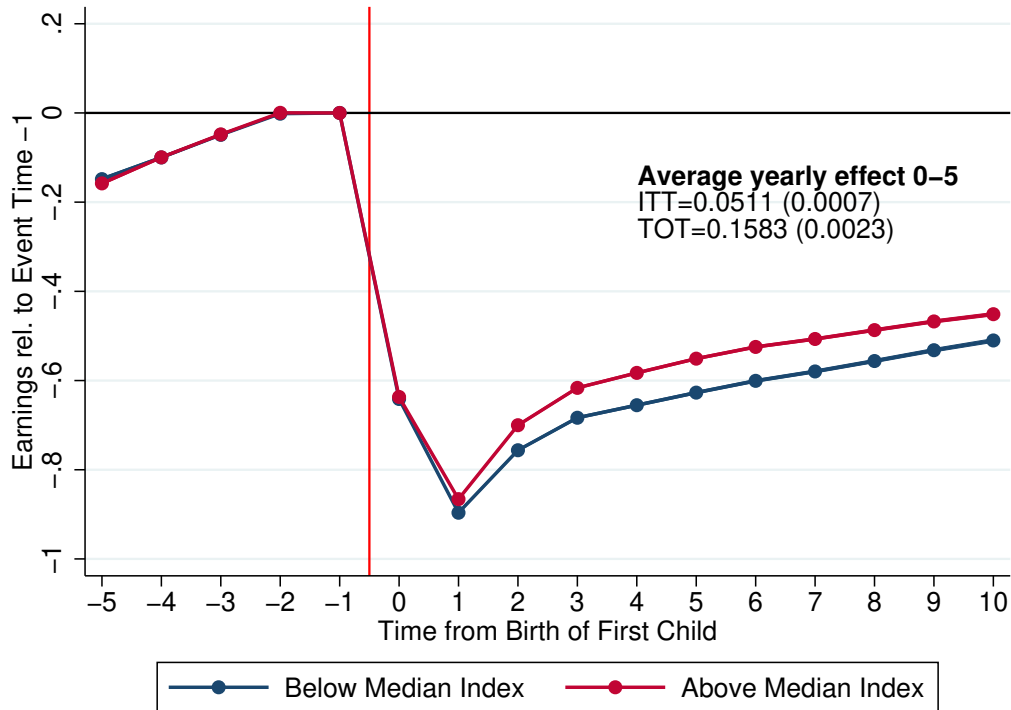
Notes: The figures shows the estimated effects of the introduction of the 1961 reform of parental leave on the dynamics of female earnings (left panel) and female employment (right panel), based on the REV sample of mothers who give birth to their first child in 1959 or 1961 and are eligible for parental leave (52 weeks of employment in the two years before birth). In each panel, we report the estimates $\hat{\alpha}_t^T$ from (4), which correspond to the percentage-point change in the child penalty at event time t for women who become eligible to 12 months of maternal leave. Shaded areas correspond to 95% confidence intervals.

Figure 5: Parental Leave Expansions: Causal Effects by Baseline Duration



Notes: This figure reports the estimated marginal effects of parental leave extensions at different levels of baseline PL duration, on the average yearly child penalty in the first five years following a first child birth. The three estimates on the right of the graph correspond to the effects of 6 months increments in parental leave duration, keeping job protection fixed at 24 months, obtained from our estimated effects of the 1990, 1996 and 2000 reforms in Figure 3, and assuming additivity in the effects of marginal extensions. The estimate on the left corresponds to the estimated effect of the 1961 introduction of parental leave, which captures the effect of a 10 month extension (from 2 to 12 months) of paid parental leave, covered with job protection.

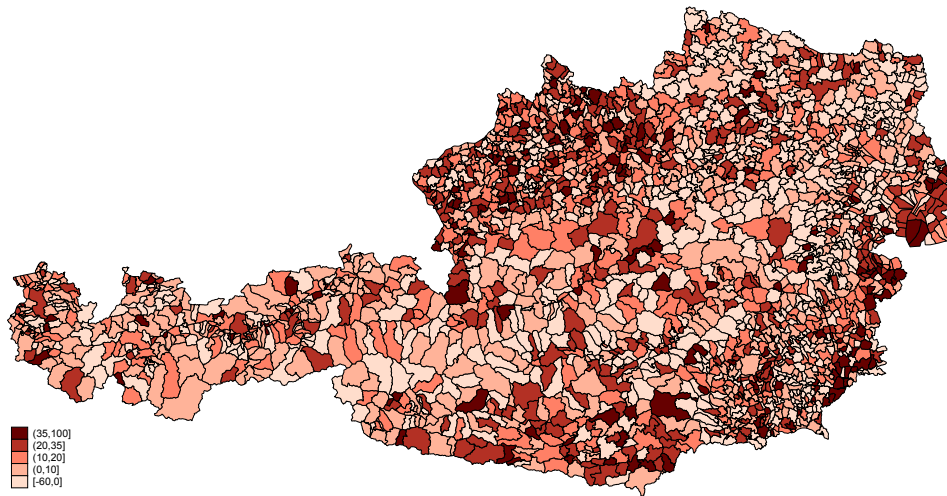
Figure 6: Child Penalty by Level of Child Care Provision
Index 1-5 in 1990



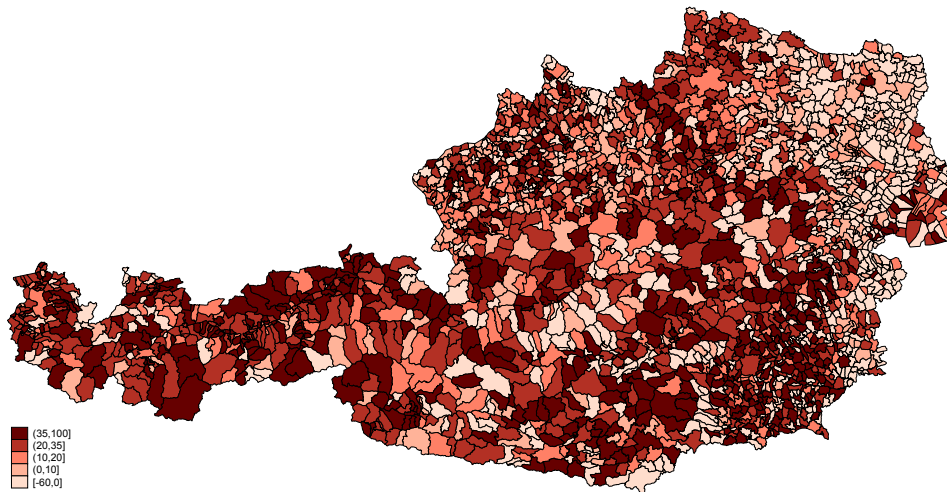
Notes: This figure compares child penalties for women who give birth in municipalities with high vs low child care coverage. We focus on mothers whose first child was born between July 1990 and June 1996. We rank municipalities according to the value of their aggregate child care index (Index 1-5) in 1990 and split the sample between municipalities below vs above the median. We then compute for each subsample the child penalty in earnings following equation 2, as done in Figure 1. To control for differences in pre-birth characteristics of mothers across the two subsamples, we use inverse-probability-weighting based on pre-birth earnings, pre-birth employment, and age. The ITT estimate corresponds to the average difference in penalties in the two groups over the first five year after birth. The TOT estimate corresponds to the ITT estimate scaled by the difference in our index 1-5 of child care coverage across the two groups.

Figure 7: Spatial Variation in Child Care Expansion

(a) Change in Index 1-5 between 1990 and 2000



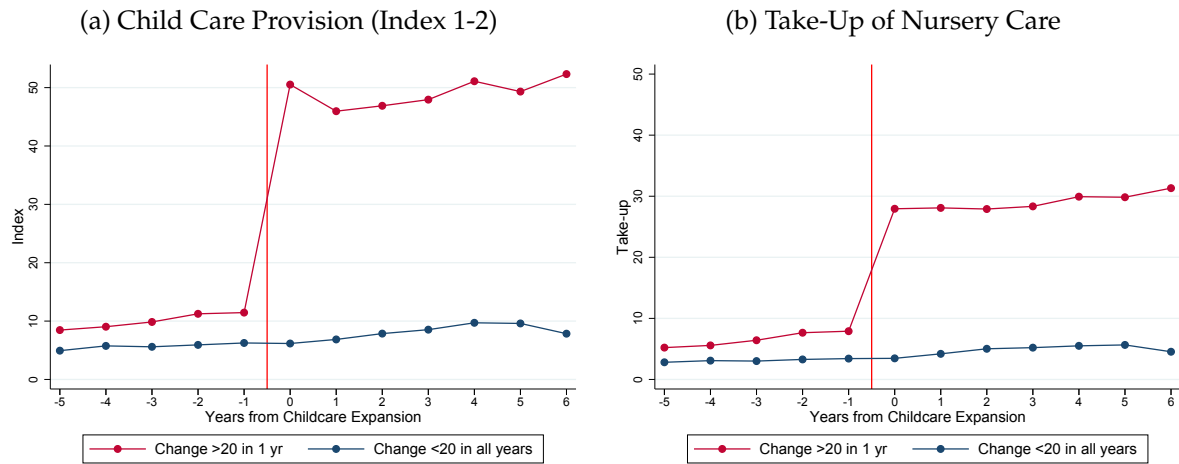
(b) Change in Index 1-5 between 2000 and 2010



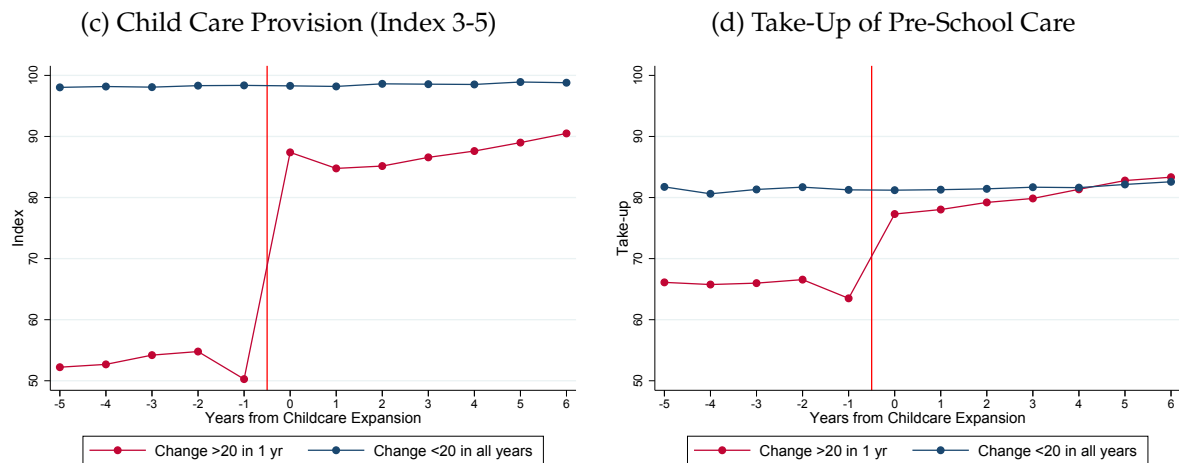
Notes: The figure showcases the significant amount of spatial variation in the expansion of local child care provision in Austria. Each panel corresponds to a heatmap that reports for each municipality the change in the value of the child care index 1-5 from 1990 to 2000 (panel (a)) and from 2000 to 2010 (panel (b)). A darker color indicates a larger increase in local child care provision. See section 4.1 for further details on the construction of the child care provision index. GIS files of municipality borders provided by Statistik Austria.

Figure 8: Local Child Care Shocks: Evolution of Child Care Provision and Take-Up for Treated and Control Municipalities

Nursery Care Expansions



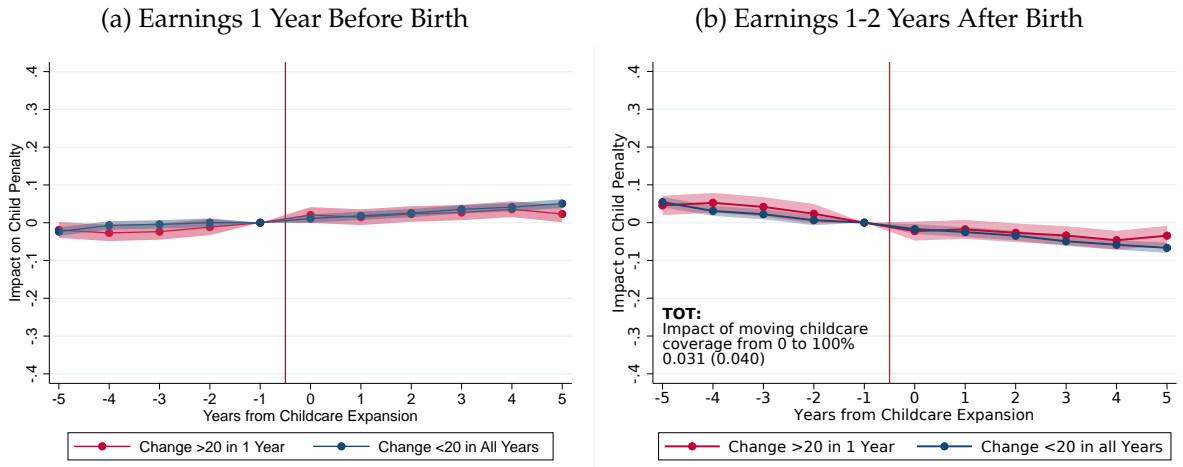
Pre-School Care Expansions



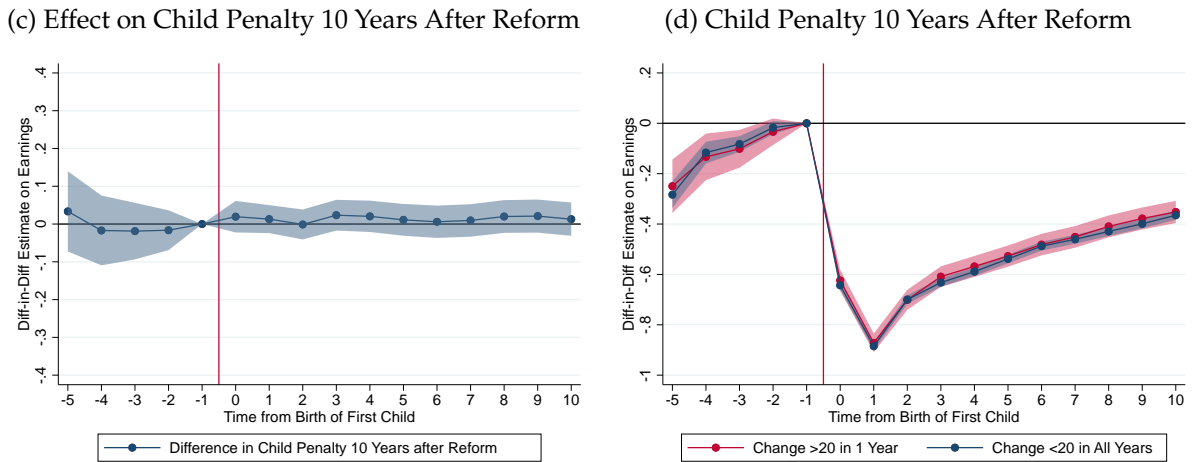
Notes: The figure shows the evolution of child care provision (left panels) and of the take up of institutional care (right panels) for treated and control municipalities around the time of a local shock to child care provision. A nursery care (resp. pre-school care) expansion event is defined as an increase of more than 20 points in the municipal nursery care (resp. pre-school care) index in one single year. Municipalities with an expansion event constitute the treated group, while municipalities that never experience an event constitute the control group. We randomly assign an expansion event year to control municipalities such that the distribution over calendar years is balanced between the two groups.

Figure 9: Estimated Effects of Nursery Care Expansions

Difference-in-Differences Evidence



Long-Run Effects

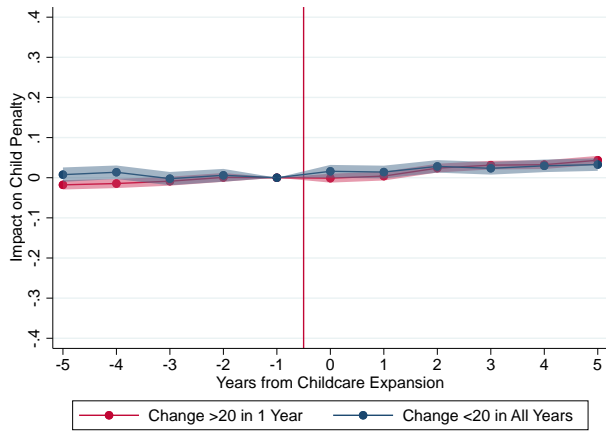


Notes: The figure shows DiD estimates of the effects of nursery care expansions on the full dynamics of female earnings. In panel (a) (resp. panel (b)), we plot the evolution of earnings at event time $t = -1$ (resp. $t \in [1, 2]$), that is for women who are one year before (resp. between one and two years after) giving birth to their first child, in treated vs control municipalities, around the event of a nursery care expansion. In panel (b) we also report the TOT effect, which corresponds to the DiD estimate \hat{a}_{st}^T for earnings scaled by the first-stage, i.e. the equivalent DiD estimate for the change in the nursery care index. In panel (c), we plot the DiD estimate \hat{a}_{st}^T for earnings at all event times t w.r.t. birth of the first child, evaluated at $s = 10$, that is 10 years after the local nursery care expansion. Panel (c) is therefore the mirror image, for nursery care reforms, of panels (a), (c) and (e) of Figure 3 in the case of parental leave reforms. Based on the estimates of panel (c), we draw in panel (d) counterfactual child penalties under the pre and post nursery care expansion regimes. Standard errors are clustered at the municipal level, and shaded areas correspond to 95% confidence intervals.

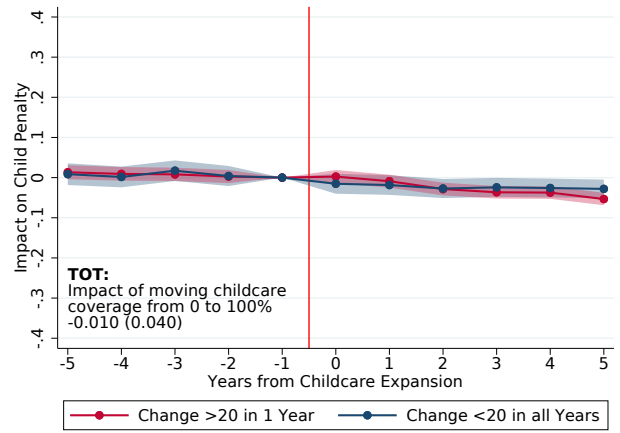
Figure 10: Estimated Effects of Pre-School Care Expansions

Difference-in-Differences Evidence

(a) Earnings 1 Year Before Birth

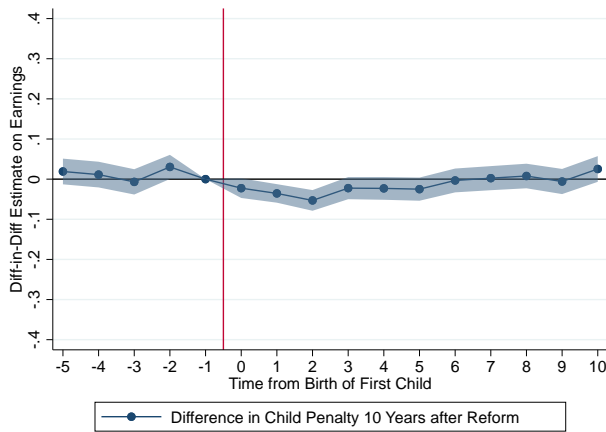


(b) Earnings 3-5 Years After Birth

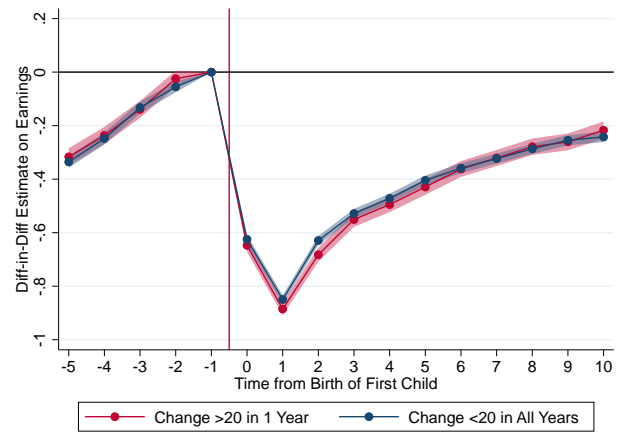


Long-Run Effects

(c) Effect on Child Penalty 10 Years After Reform

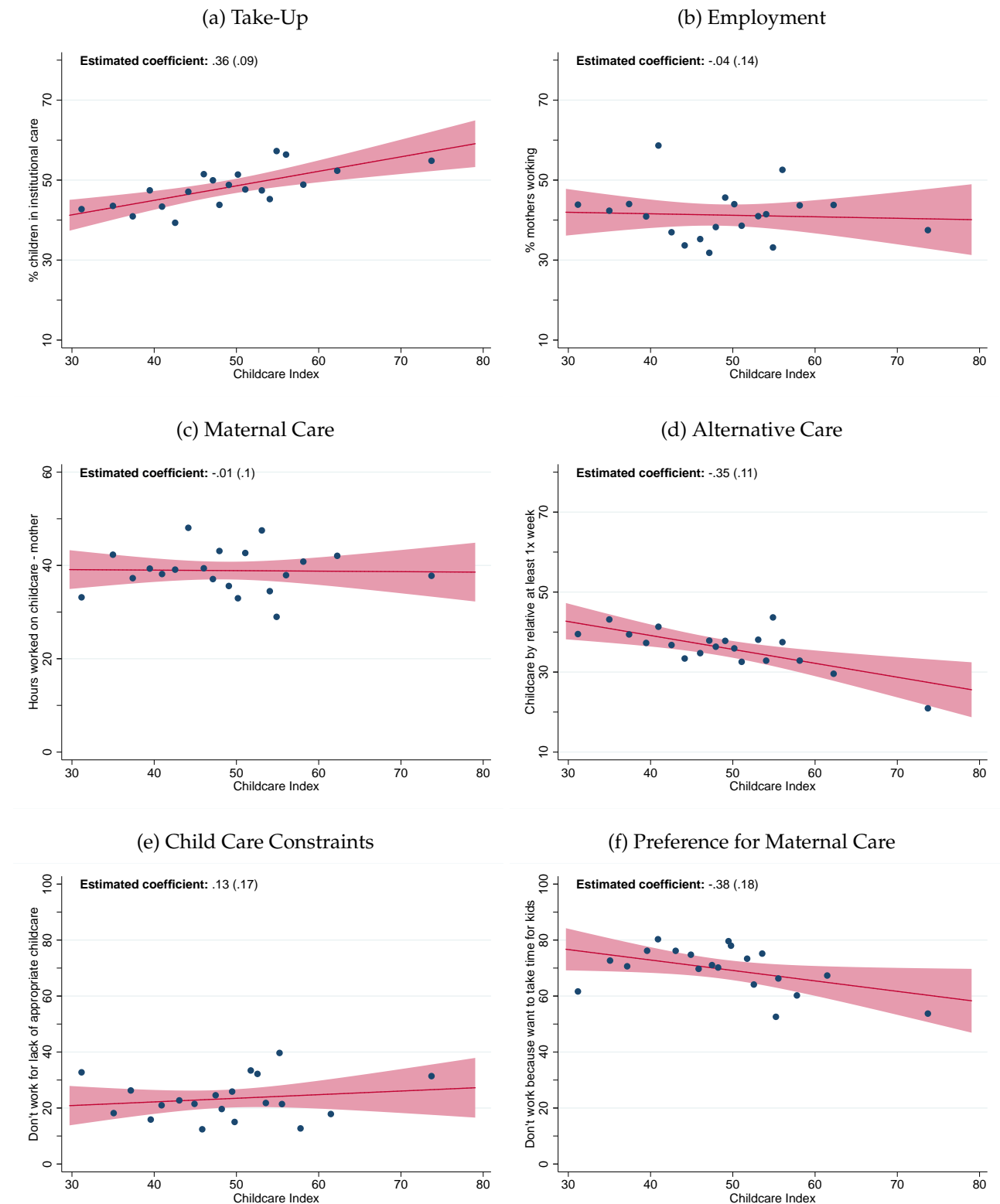


(d) Child Penalty 10 Years After Reform



Notes: The figure replicates the analysis of Figure 9 for pre-school care expansions events. It shows the DiD estimates of the effects of pre-school care expansions on the full dynamics of female earnings. In panel (a) (resp. panel (b)), we plot the evolution of earnings at event time $t = -1$ (resp. $t \in [1, 2]$), that is for women who are one year before (resp. between one and two years after) giving birth to their first child, in treated vs control municipalities, around the event of a pre-school care expansion. In panel (b) we also report the TOT effect, which corresponds to the DiD estimate $\hat{\alpha}_{st}^T$ for earnings scaled by the first-stage, i.e. the equivalent DiD estimate for the change in the pre-school care index. In panel (c), we plot the DiD estimate $\hat{\alpha}_{st}^T$ for earnings at all event times t w.r.t. birth of the first child, evaluated at $s = 10$, that is 10 years after the local pre-school care expansion. Panel (c) is therefore the mirror image, for pre-school care reforms, of panels (a), (c) and (e) of Figure 3 in the case of parental leave reforms. Based on the estimates of panel (c), we draw in panel (d) counterfactual child penalties under the pre and post pre-school care expansion regimes. Standard errors are clustered at the municipal level, and shaded areas correspond to 95% confidence intervals.

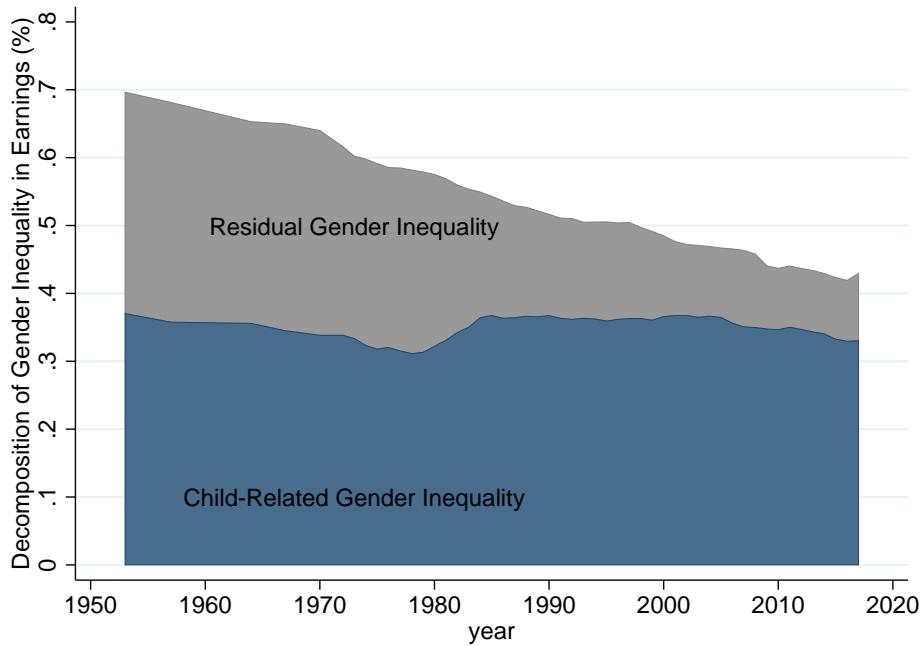
Figure 11: Reported Child Care Arrangements and Preferences: Residual Correlations With Child Care Provision



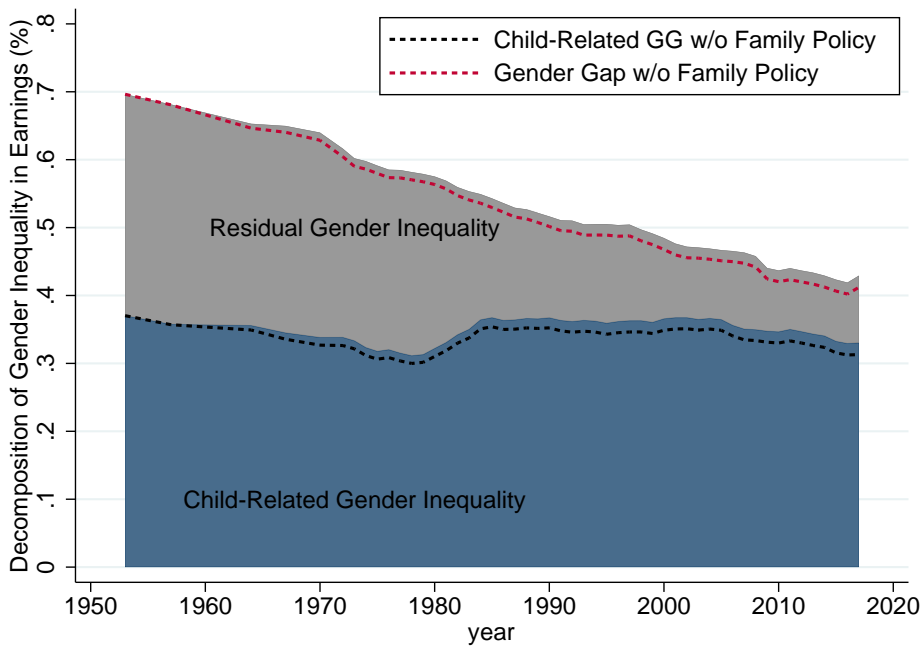
Notes: The figure reports correlations between measures of institutional child care provision at the political district level and household level information on time use, preferences and child care arrangements from two large external surveys from the Austrian Microcensus in 1995 and 2002. We pool observations from both waves of the survey. Each panel corresponds to a different outcome. We control for differences in the education and age structure across political districts by regressing the average outcome in a political district on the fraction of the population in three education bins and five age bins. Each panel plots a binscatter of the residuals, rescaled at the average level of education and age in Austria, against the value of the child care index (index 1-5). We also report on each panel the residual correlation estimate and its standard error. Panel (a) focuses on the fraction of children aged 2 to 5 in institutional child care. Panel (b) looks at the fraction of women who report working in the survey. Panel (c) focuses on the average number of hours that mothers report spending weekly on child care, while panel (d) looks at the fraction of children that are under the care of relatives. Panel (e) uses as an outcome the fraction of mothers reporting that their career is constrained due to a lack of appropriate child care, while the outcome in panel (f) is the fraction of women who report that they do not work because they do want to take care of their children.

Figure 12: Gender Gap Decomposition & Counterfactual Inequality Absent Family Policies

(a) Decomposition of Gender Gap Between Child-Related and Residual Inequality



(b) Gender Gap Estimates Absent Any Family Policy



Notes: This figure decomposes the gender gap in Austria for the period 1953 to 2017, and provides counterfactual gender inequality estimates absent family policies. Panel (a) plots the results of the decomposition of the earnings gap (corresponding to the sum of the blue and grey shaded areas) into child-related (blue) and residual factors (grey), following the methodology of [Kleven, Landais and Sogaard \(2019\)](#). In Panel (b), we superimpose the result of a counterfactual decomposition where we undo the effects of all family policy reforms since 1953, that is, we set parental leave duration and our child care indices to zero in all years. This decomposition is based on the reduced form estimates of the dynamic effects of all family policy reforms, from sections 3 and 4. See text for details.

Table 1: Gender Gap Decomposition Under Observed & Counterfactual Family Policy Regimes

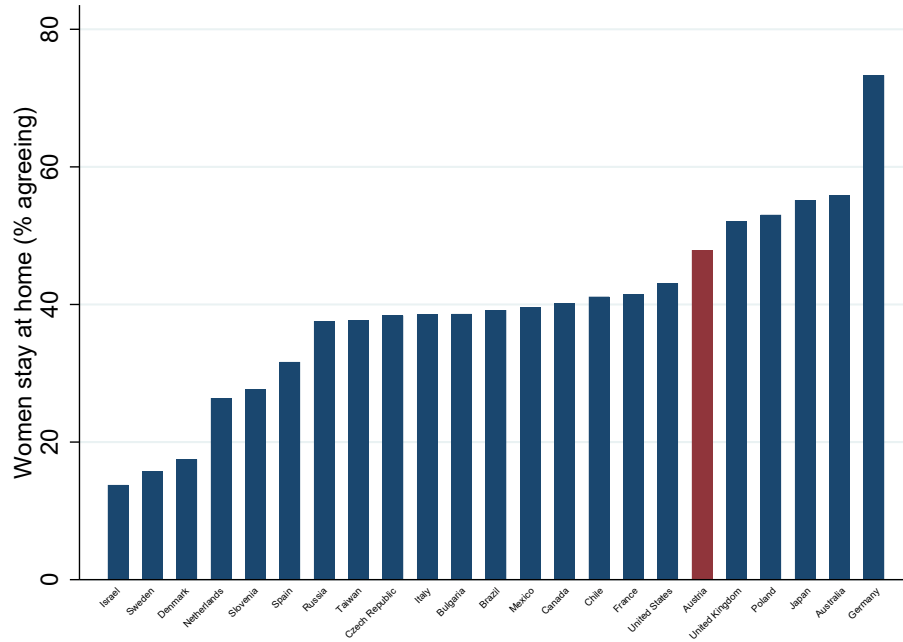
Year	1953	1964	1970	1980	1990	2000	2010
A. Gender Inequality Decomposition							
Share Childless	0.217	0.180	0.163	0.151	0.151	0.157	0.167
Share Mothers t<0	0.144	0.151	0.153	0.186	0.186	0.177	0.208
Share Mothers t=0-10	0.172	0.173	0.184	0.197	0.217	0.208	0.188
Share Mothers t=11-20	0.166	0.160	0.160	0.163	0.169	0.187	0.187
Share Mothers t=21-30	0.166	0.168	0.166	0.153	0.151	0.156	0.160
Share Mothers t>30	0.134	0.168	0.173	0.151	0.125	0.113	0.083
Earnings Women	307	790	1,308	3,877	6,540	9,705	13,095
Earnings Men	1,011	2,277	3,634	9,128	13,526	18,833	23,258
Employment Rate Women	0.495	0.515	0.486	0.522	0.538	0.599	0.646
Employment Rate Men	0.904	0.887	0.858	0.826	0.775	0.759	0.744
<i>Decomposition Btw Child-Related & Residual Inequality</i>							
Earnings Gap	0.696	0.653	0.640	0.575	0.517	0.485	0.437
Child-Related Gender Gap	0.370	0.356	0.338	0.322	0.367	0.366	0.347
B. Counterfactual Family Policy Scenario							
<i>No Parental Leave and No Child Care (Baseline)</i>							
Earnings Gap	0.696	0.646	0.628	0.562	0.499	0.463	0.414
Child-Related Gender Gap	0.370	0.342	0.314	0.296	0.331	0.323	0.301
<i>No Parental Leave and No Child Care (Optimistic)</i>							
Earnings Gap	0.696	0.648	0.632	0.565	0.503	0.474	0.428
Child-Related Gender Gap	0.370	0.347	0.322	0.302	0.340	0.345	0.329
<i>No Parental Leave and No Child Care (Pessimistic)</i>							
Earnings Gap	0.696	0.644	0.624	0.559	0.495	0.452	0.400
Child-Related Gender Gap	0.370	0.337	0.306	0.289	0.323	0.301	0.272

Notes: Panel A reports inputs and results of our decomposition exercise of gender gaps for selected years. In particular, we start by reporting the fraction of women ψ_{st} who are at each event time t relative to the birth of their first child in year s . We also report average nominal earnings of women and men in each year in Euro (we use the ECB reference rate to convert Austrian Shillings to Euro for years before the introduction of the Euro) estimated from the ASSD earnings distribution (employed only). The employment rates of women and men are obtained from the census. Finally, we report the actual and child-related earnings gaps corresponding to Figure 12 panel (a). In Panel B we report the results of our counterfactual decomposition exercise, where we undo the effects of all family policies since 1953. We first report the results from our baseline scenario, and then the results from two bounding exercises. The first, “optimistic” exercise, uses, instead of our point estimates, the upper bound of the 95% confidence interval on $\frac{dP_t}{d\tau_k}(\tau_k)$ for each policy change. The second, or “pessimistic” scenario, uses the lower bound of the 95% confidence interval. See text for details.

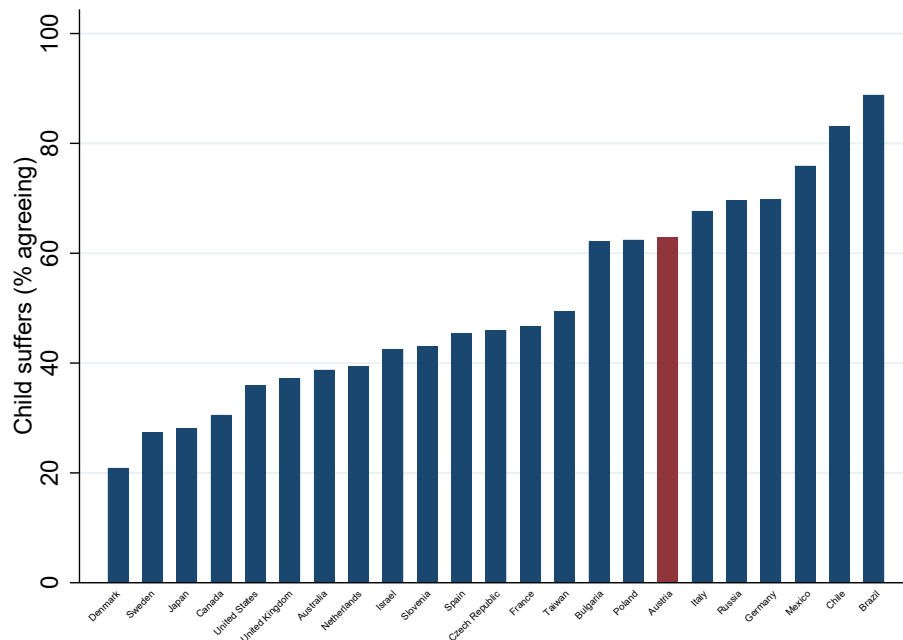
A Supplementary Figures: Context

Figure A.I: Agreement with Statements

(a) "A Woman Should Stay Home when She Has a Child under School Age"



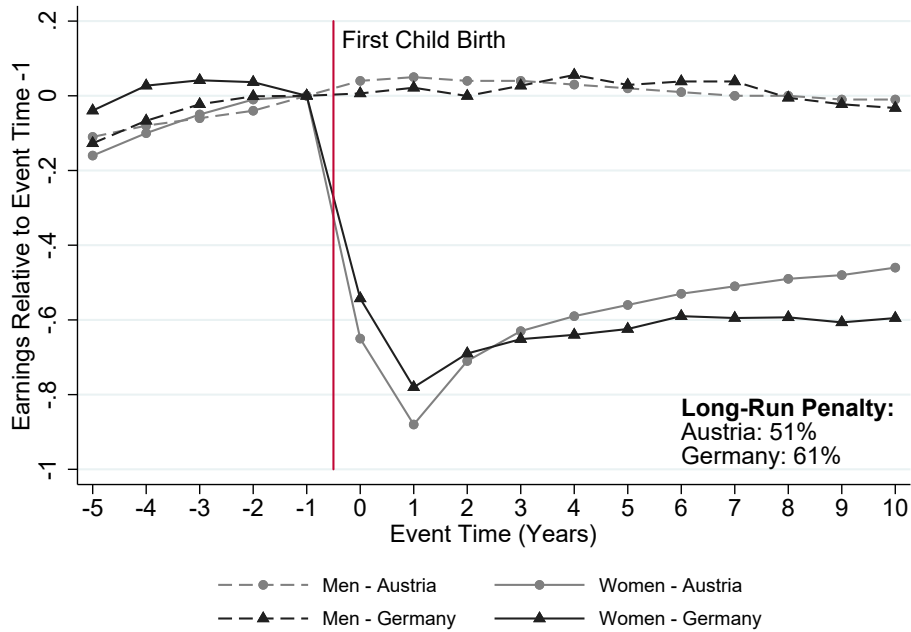
(b) "When a Mother Works for Pay, Her Children Suffer"



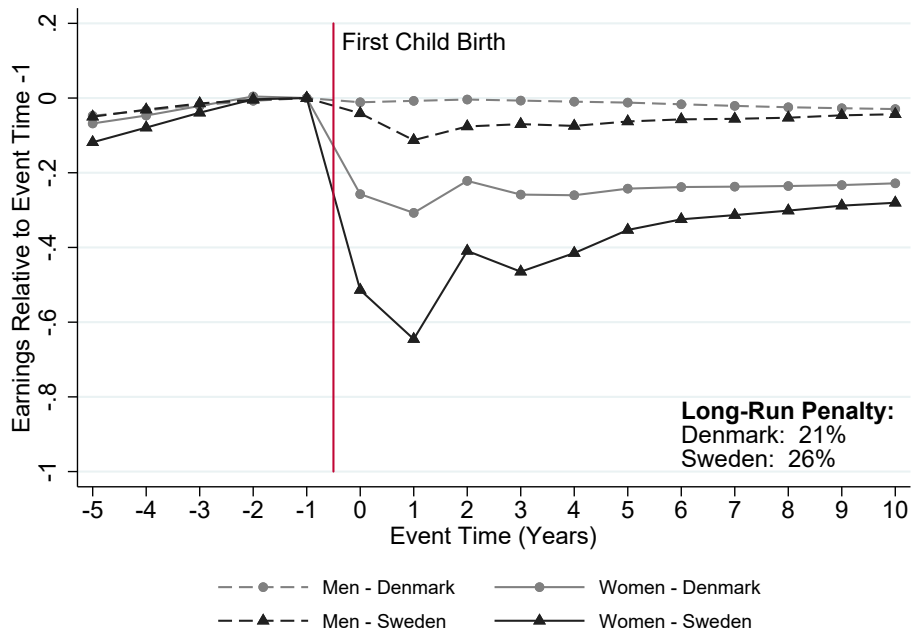
Notes: The figure is based on data from the 2012 wave of the International Social Survey Program (ISSP).

Figure A.II: Child Penalties - Other Countries

(a) German-Speaking Countries



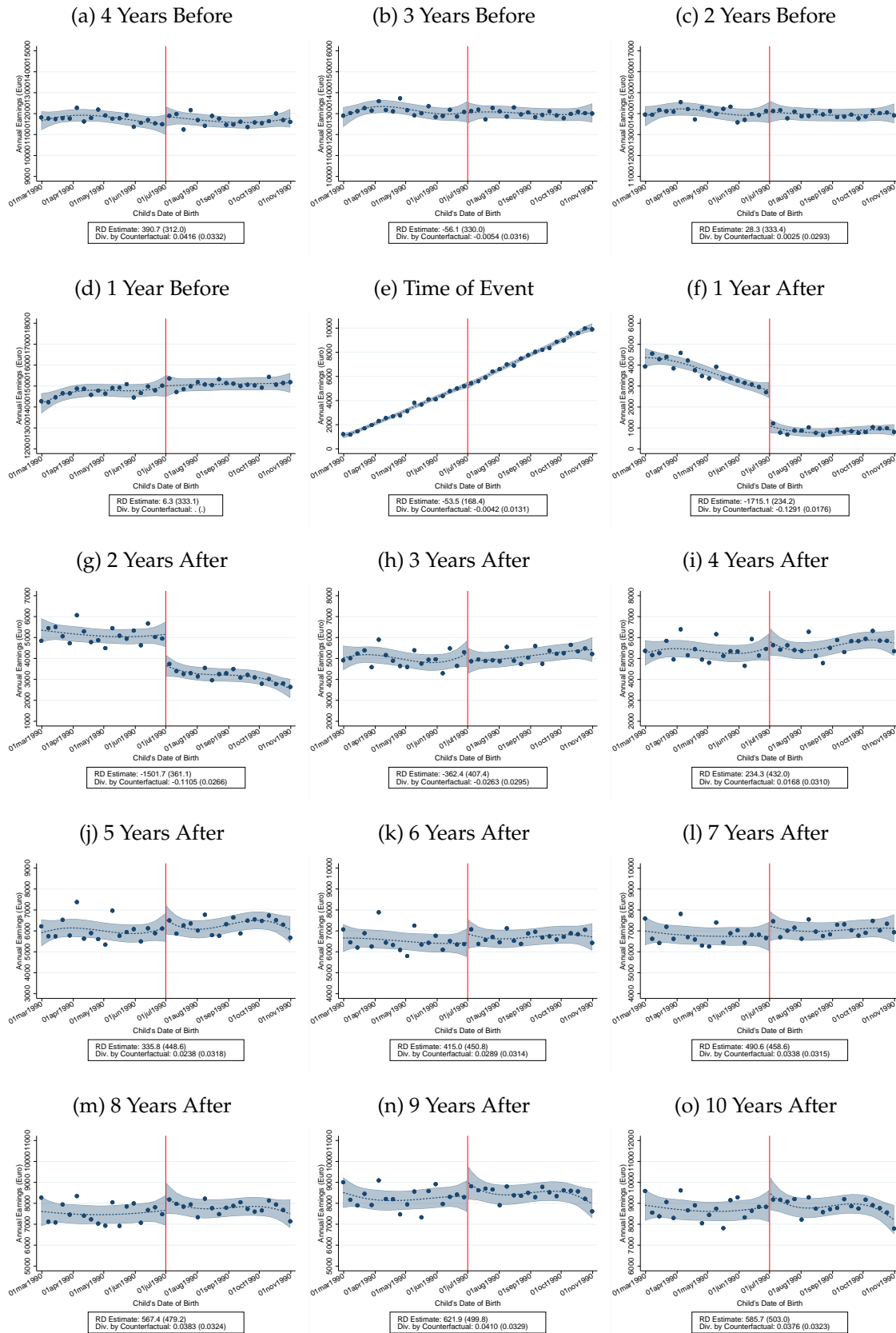
(b) Scandinavian Countries



Notes: Figures reproduced from [Kleven, Landais, Posch, Steinhauer and Zweimuller \(2019\)](#). We plot child penalties P_t^g as defined in equation (2) for each gender g by event time t (years from birth of first child) for a set of countries. Annual earnings are zero if an individual is not working in a given year. Long-run penalties are defined as $P_{10}^m - P_{10}^w$. For details on data sources and methodology see [Kleven, Landais, Posch, Steinhauer and Zweimuller \(2019\)](#).

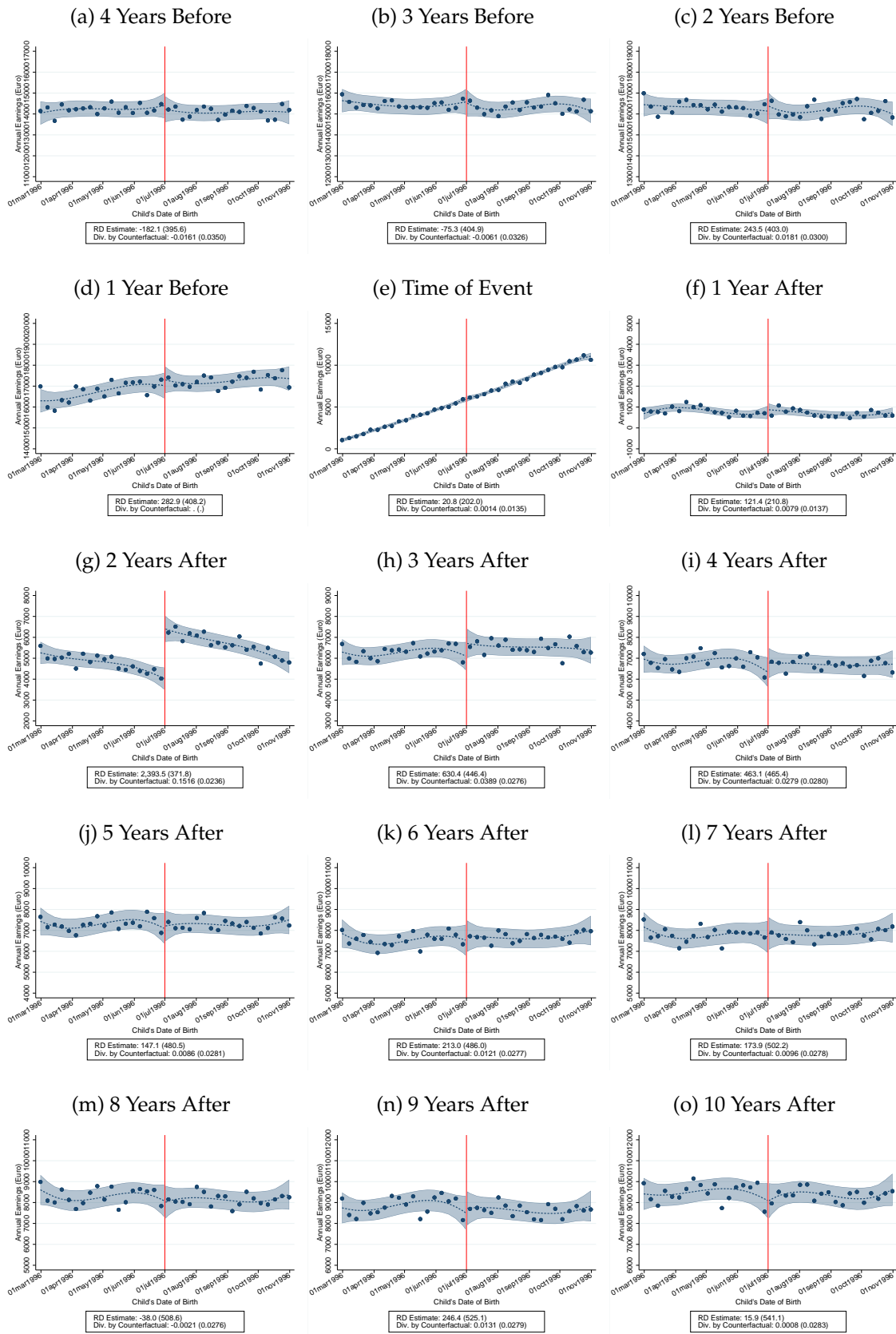
B Supplementary Figures: Parental Leave

Figure B.I: RDD Evidence on Maternal Earnings: 1990 Parental Leave Reform



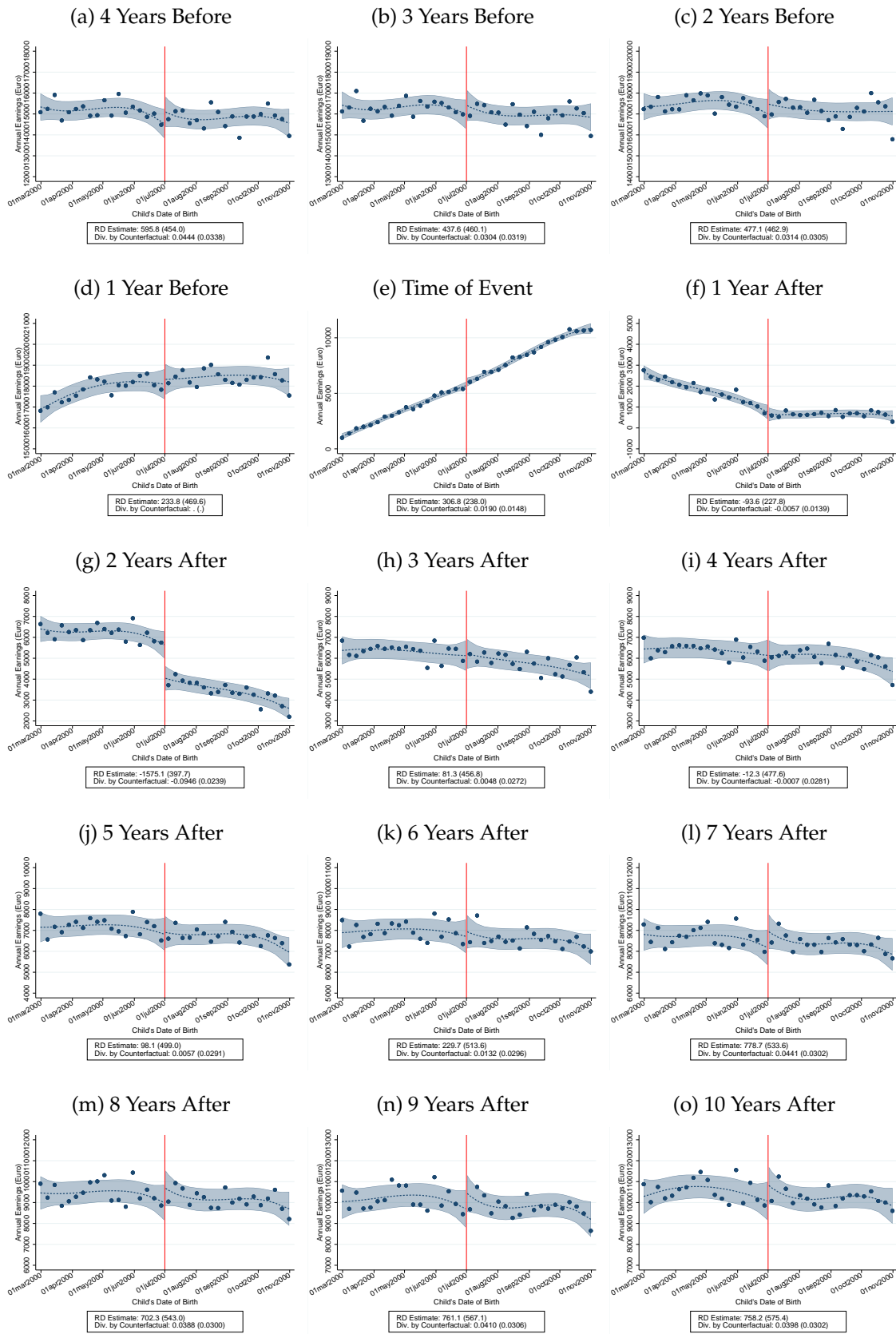
Notes: Figures plot the average annual earnings of mothers by week of first birth for each event time relative to first birth. We adjust earnings by CPI (base year 2000) and correct for topcoding (see Appendix D.1). Earnings are zero if the individual did not work in a particular year. We include first births in a 4-month window around the reform cut-off dates (July 1st 1990, 1996, and 2000). We exclude women ineligible for parental leave (less than 52 weeks of work in the two years preceding birth). For each event time separately we regress annual earnings on age dummies and a cubic in distance to the cut-off, allowing for separate trends on each side. We plot cubic trends from this regression including 95% confidence intervals. We report the discontinuity estimate and standard error as well as the estimate scaled by average counterfactual earnings at the same event time from child penalty regressions estimated on pre-reform samples.

Figure B.II: RDD Evidence on Maternal Earnings: 1996 Parental Leave Reform



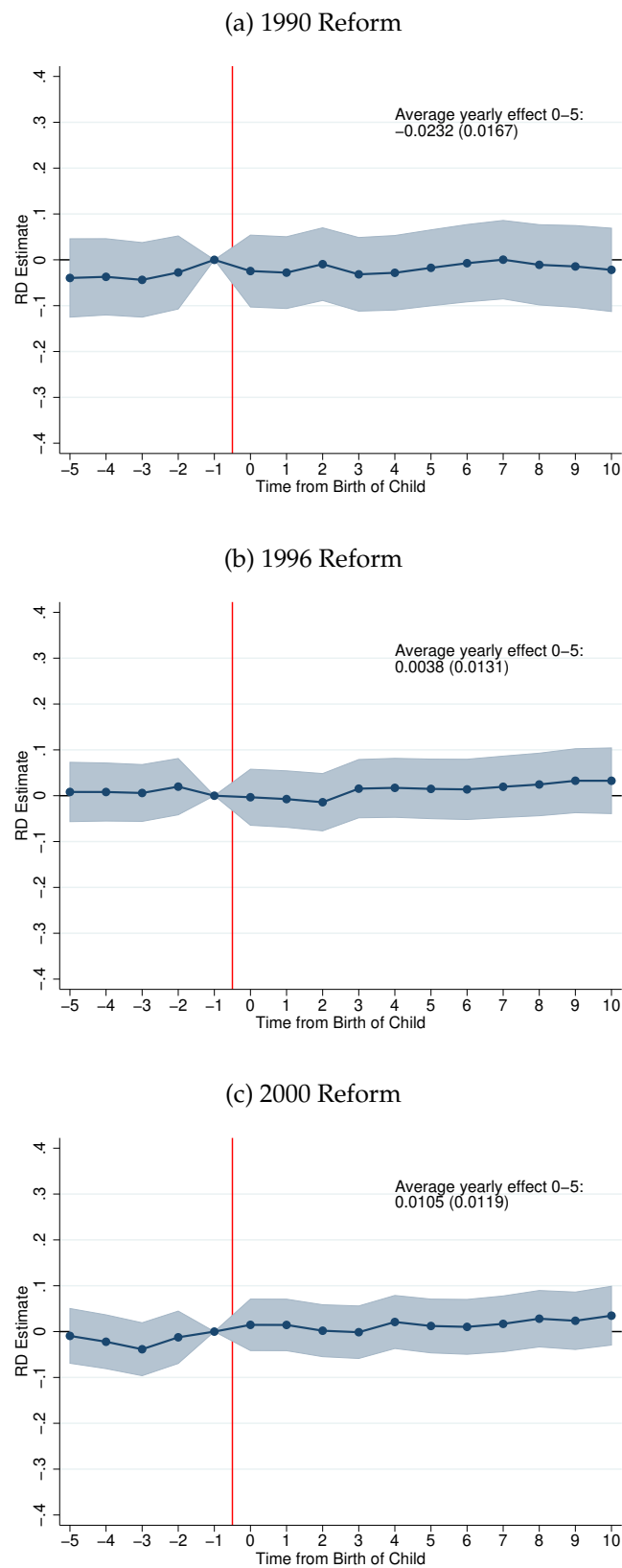
Notes: See notes to Figure B.I.

Figure B.III: RDD Evidence on Maternal Earnings: 2000 Parental Leave Reform



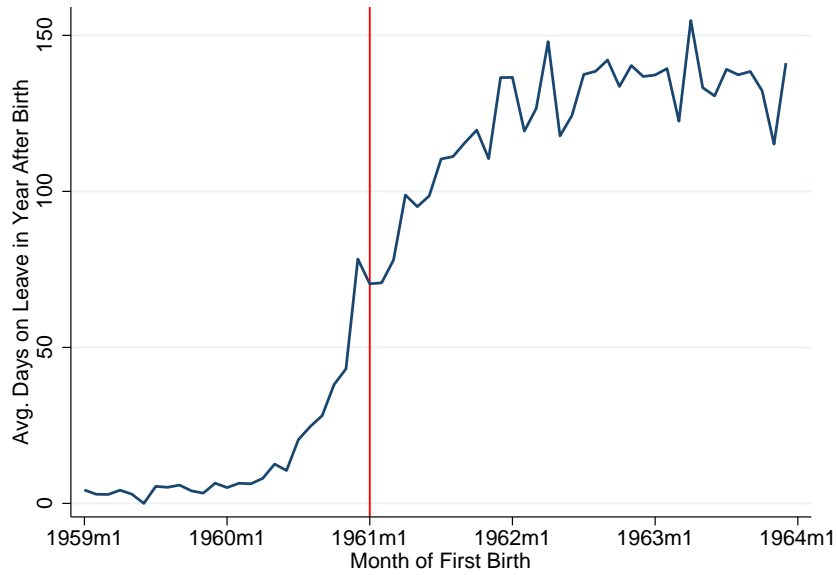
Notes: See notes to Figure B.I.

Figure B.IV: Effect on Paternal Earnings: 1990, 1996 and 2000 Parental Leave Reforms



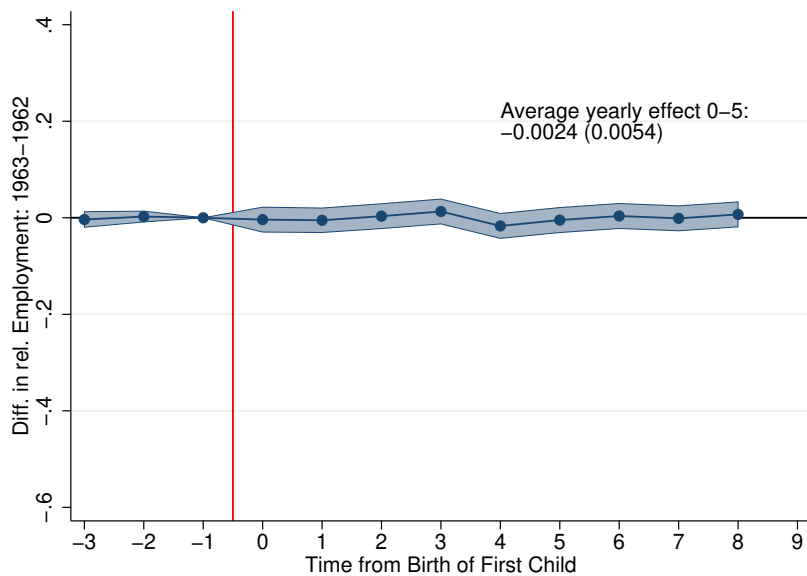
Notes: The figure replicates Figure 3, but on the sample of first time fathers. It displays estimates of the effects of three parental leave reforms on the dynamic earnings of fathers. Fathers are matched to first births based on administrative data on applications for child benefits. See notes to Figure 3 for further details.

Figure B.V: 1961 Reform: Take-Up



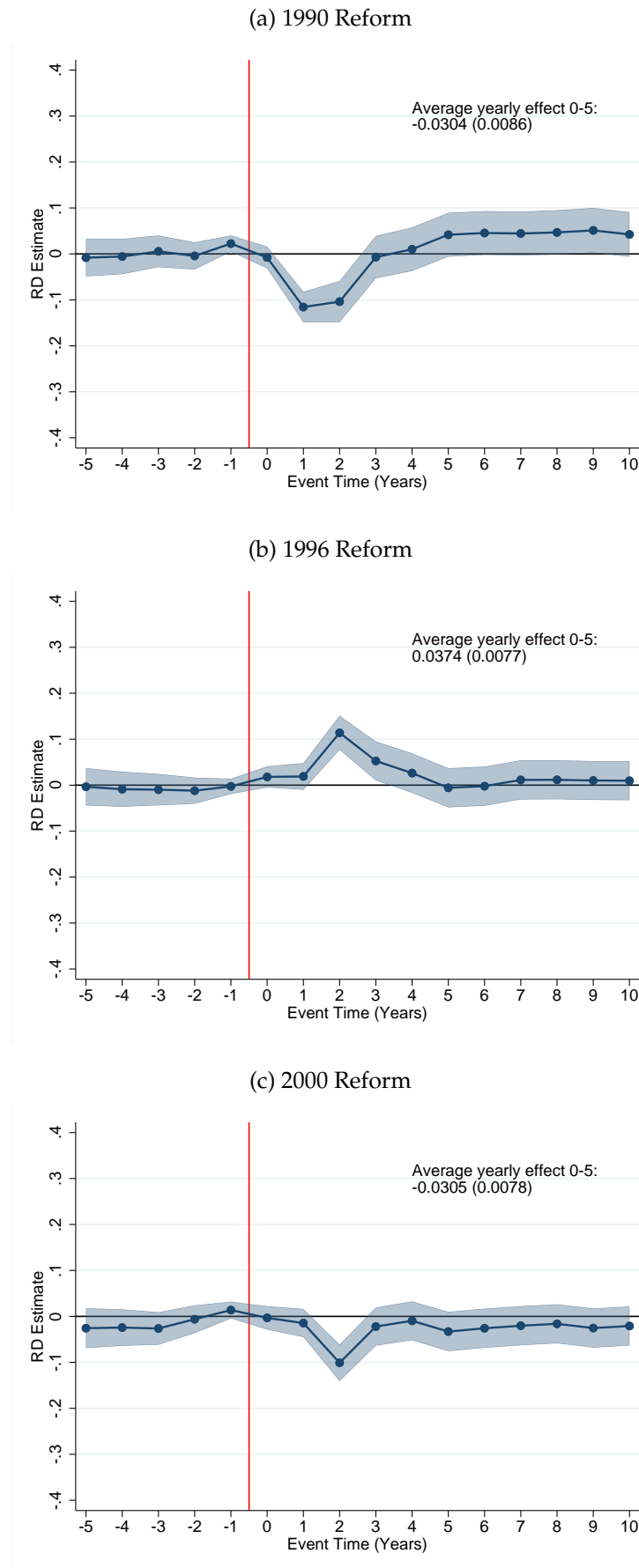
Notes: The figure shows the evolution of maternity leave take up around the 1961 reform which introduced parental leave. We plot average days on parental leave in the year following birth of the first child, by month of first birth (based on REV data). As the reform was grandfathered, women who gave birth in 1960 became eligible, hence the observed gradual increase in take up.

Figure B.VI: 1961 Reform: Robustness to Trends



Notes: The figure investigates the presence of trends that may confound the estimated effects of the 1961 reform. The graph shows placebo results from specification 4 estimated on a sample of mothers who gave birth in 1962 ($T = 0$) or 1963 ($T = 1$). See notes to Figure 4.

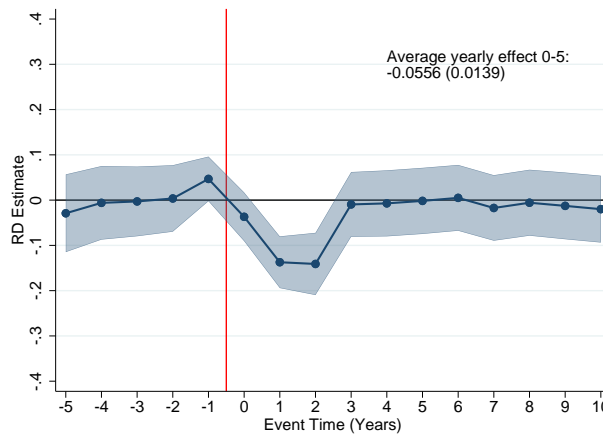
Figure B.VII: Earnings Effect of Parental Leave Reforms on Mothers in Top Quartile of Pre-Birth Earnings Distribution



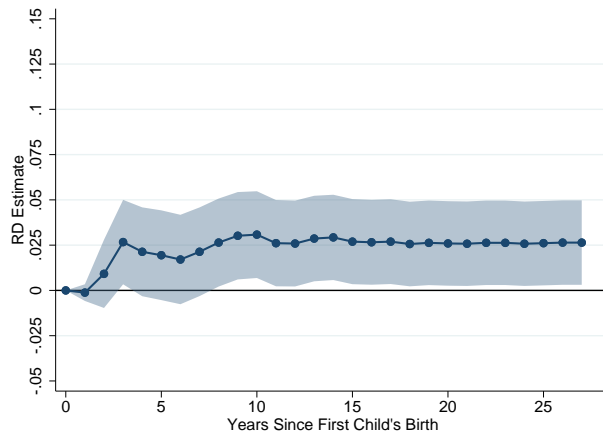
Notes: This figure investigates the presence of heterogeneous responses to parental leave reforms by mothers' earnings level. We reproduce estimates from Figure 3 on women at the top end of the earnings distribution. To this effect, we split our reform samples into four quartiles by annual earnings one year before birth. We then run exactly the same procedure as in Figure 3 keeping only mothers in the top pre-birth earnings quartile.

Figure B.VIII: Effects of Parental Leave: Mediating Role of Fertility

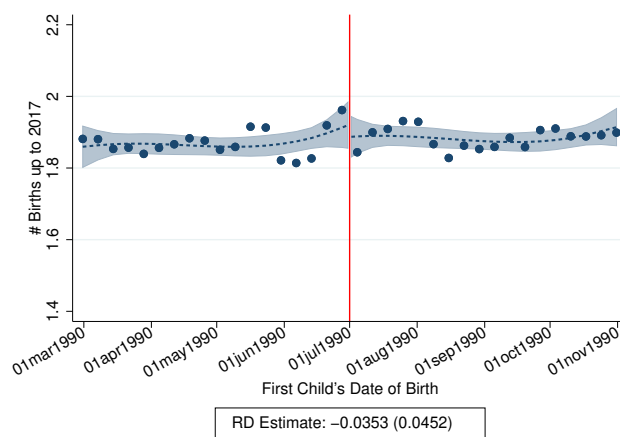
(a) Dynamic Effects of the 1990 Reform - 1 Child Only



(b) 1990 Reform - Probability of Second Birth

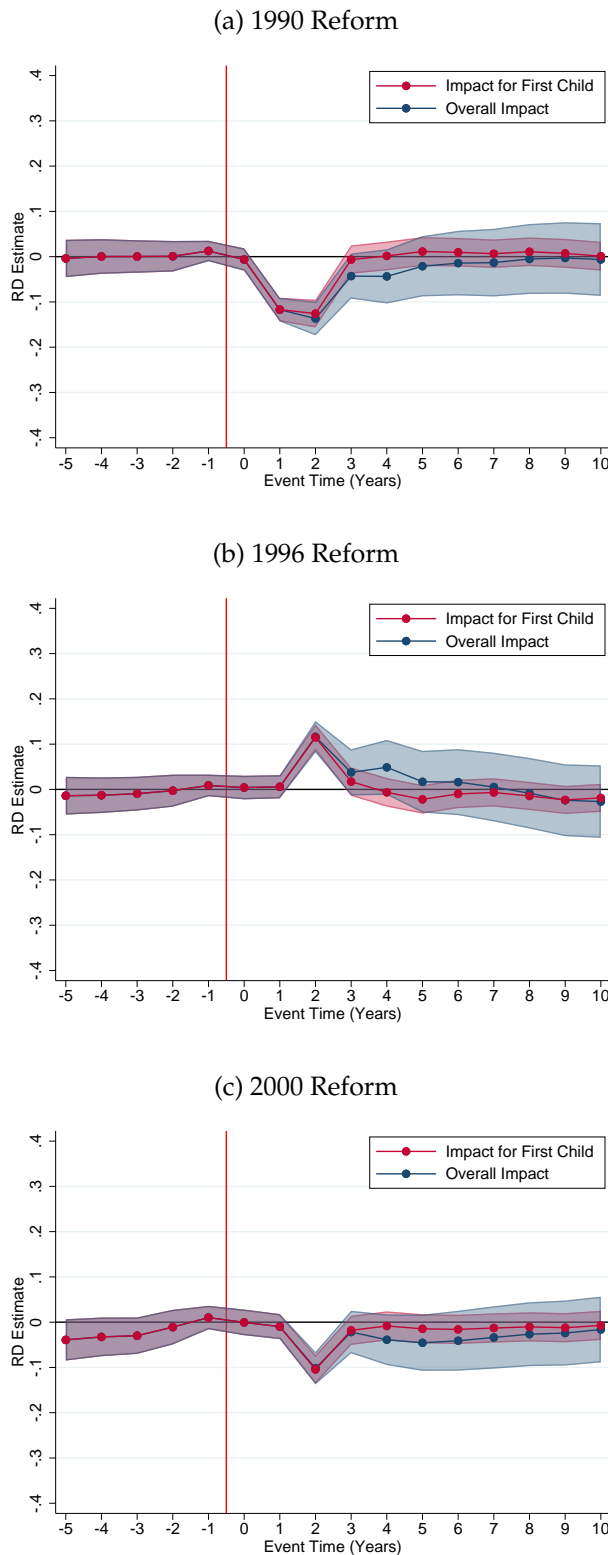


(c) Effect of the 1990 Reform on Completed Fertility



Notes: Panel (a) plots the 1990 reform effect (see notes to Figure 3) for the subsample of mothers who only have one child in the 1989-2017 period. Panel (b) plots the 1990 reform effect of the probability of a second birth by years since the first child's birth. Panel (c) plots average number of births by week of birth of the first child for mothers who gave birth to their first child between March-November 1990. Average births includes all observed births up to 2017. Dashed lines are fitted values based on a cubic polynomial separately fitted at each side of the cut off with 95% confidence interval.

Figure B.IX: Effects of Parental Leave on Child Penalties Accounting For Subsequent Births

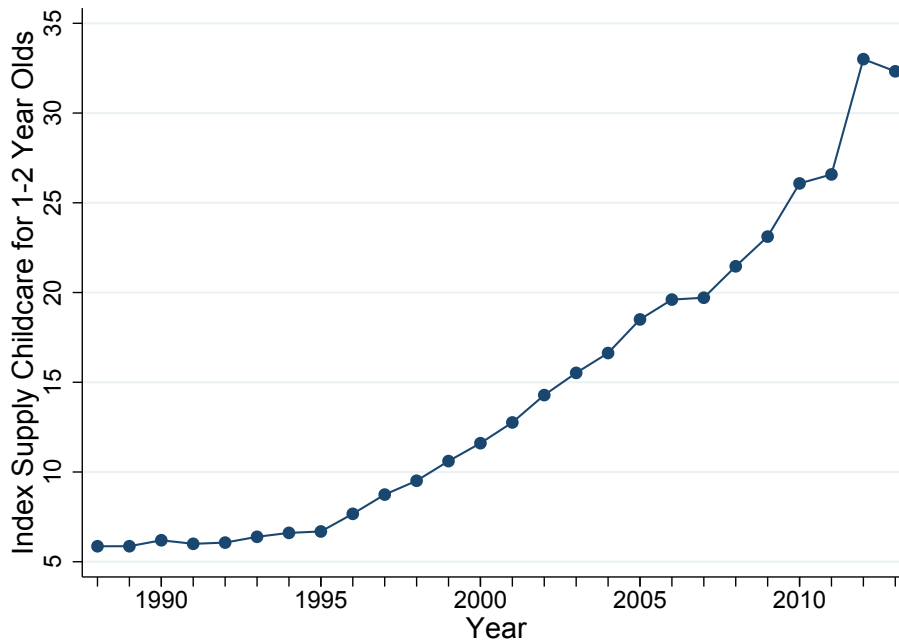


Notes: Our baseline estimates in Figure 3 identify how child penalties respond to the PL regime applicable to the first child. To compute the total effect on child penalties of moving, in steady state, from one PL regime to another, this figure accounts for the effect of PL extensions for births of higher parity. To this effect, we turn to mothers giving birth to their second, third and fourth child, and replicate, for each parity, our dynamic estimates of the impact of PL reforms on earnings. We then construct the total steady state impact of a reform by adding dynamic estimates at all parities, weighted by the average completed fertility, and average timing between parities, of women who had their first birth around the time of each reform.

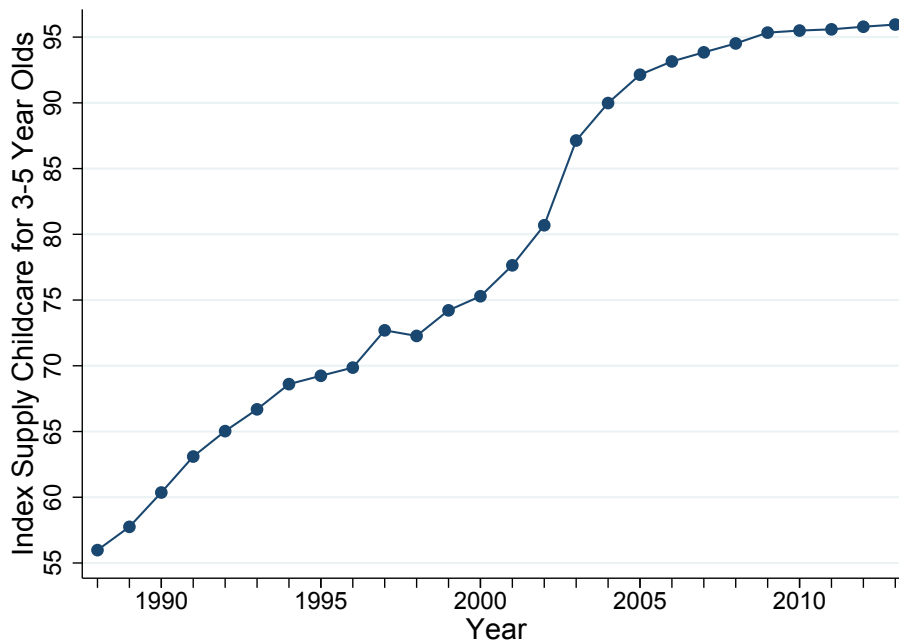
C Supplementary Figures: Child Care Provision

Figure C.I: Index of Child Care Provision over Time

(a) Child Care Index 1-2

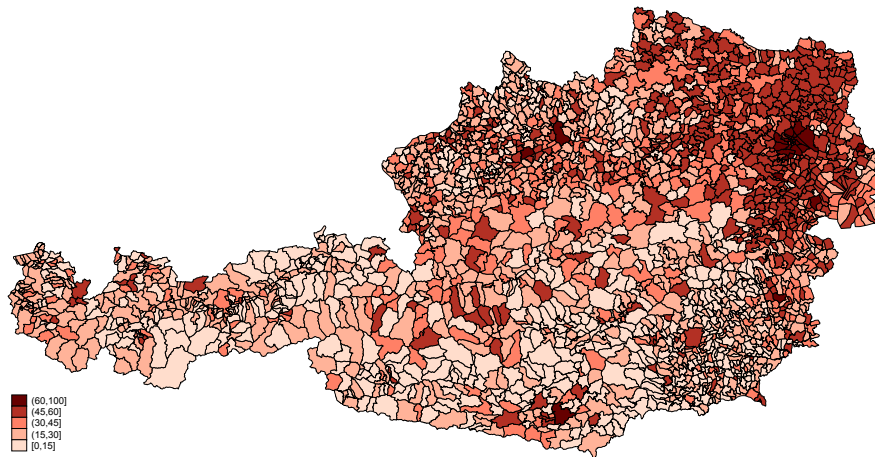


(b) Child Care Index 3-5



Notes: This figure plots our child care indices over time. Each index is calculated at the municipality level based on child care availability and the number of children in the municipality. We then calculate weighted averages by year weighting by the total number of children age 1-2 (panel a) and children age 3-5 (panel b) in the municipality.

Figure C.II: Spatial Variation in Child Care Provision - Index 1-5 - 1990



Notes: This figure plots our child care index 1-5 by municipality using Austrian GIS maps with municipality borders (administrative boundaries). We shade each municipality according to the level of the index. Darker colors indicate a higher index in 1990.

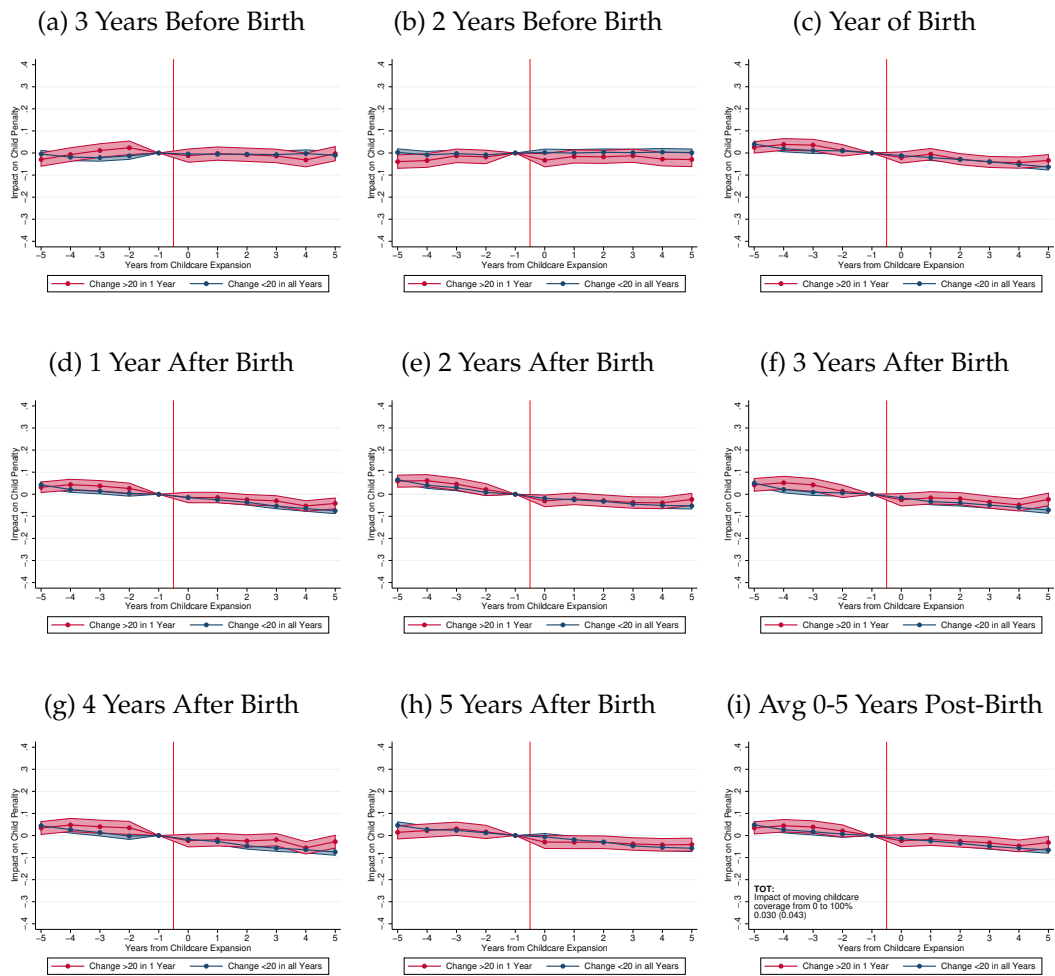
Table C.I: Descriptive Statistics of Municipalities By Treatment Status

Treatment Status	Index 1-2		Index 3-5		All
	Control	Treatment	Control	Treatment	
Panel A. 1991 Census Characteristics					
Number of Municipalities	1,554	284	253	1,507	1,887
Total Population	4,797,414	1,323,339	2,938,329	3,817,840	7,795,786
% Urban	0.39	2.11	9.49	0.07	1.48
% Towns/Suburbs	15.51	32.04	28.46	15.59	18.76
% Rural	84.11	65.85	62.06	84.34	79.76
Average Population	3,087	4,660	11,614	2,533	4,131
Empl. Rate Men	58.63	58.18	57.91	58.68	58.53
Empl. Rate Women	35.18	35.10	37.54	34.88	35.25
Female Earnings at $t=-1$	17,317	17,300	17,896	17,228	17,308
Panel B. Child Care in 1991					
% Municip. with Nursery	0.77	4.93	12.65	0.46	2.33
% Municip. with All-Day Pre-School	40.93	40.85	83.79	33.84	41.92
% of Children (1-2) in Nursery	0.00	0.01	0.02	0.00	0.00
% of Children (3-5) in Pre-School	0.57	0.59	0.82	0.53	0.57
Index 1-2	0.16	0.99	2.95	0.08	0.48
Index 3-5	50.54	49.97	85.86	45.25	51.01
Index 1-5	30.39	30.38	52.69	27.18	30.80

Notes: Panel A shows municipality characteristics in our treatment and control municipalities drawn from the 1991 census and earnings information from the ASSD. For earnings, we focus on women just about to give birth (i.e. in event time -1 with respect to the birth of their first child).

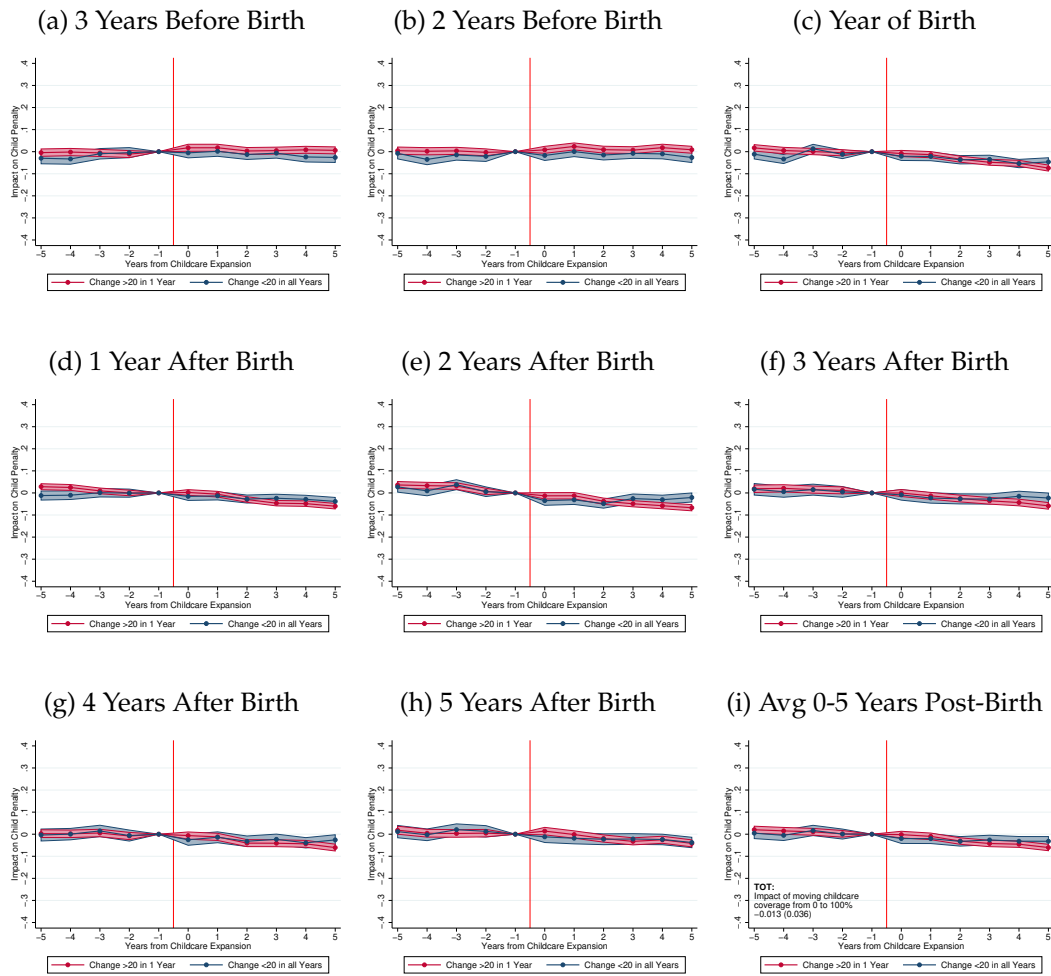
Panel B shows average child care availability according to the child care availability dataset provided by Statistics Austria. We calculate the average share with a nursery (children age 1-2), with all-day pre-school (children age 3-5) and the share of children enrolled in each of these types of child care facilities. We also show the averages of our child care indices age 1-2, 3-5 and 1-5.

Figure C.III: Nursery Care



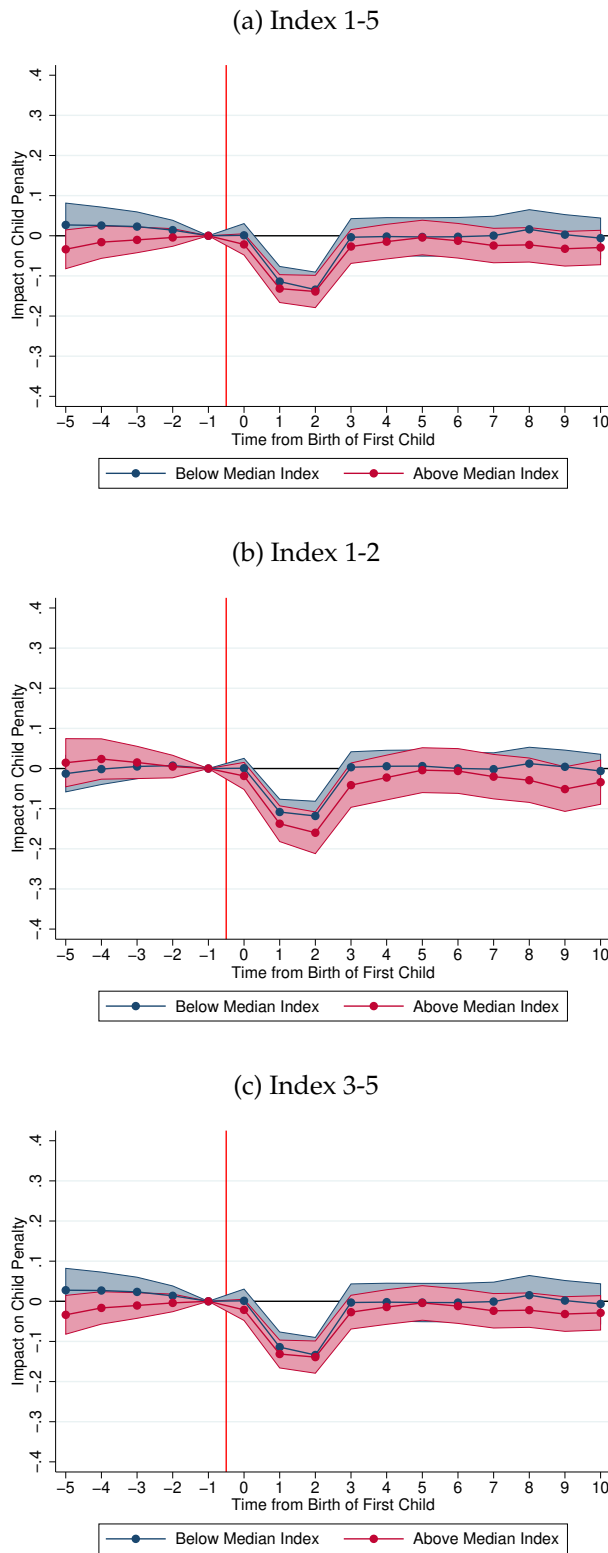
Notes: The figure plots difference-in-differences estimates from specification 6 for the effects of nursery care expansions on the full dynamics of female earnings. Each panel corresponds to different event time with respect to the birth of the first child. In panel (a) we plot the evolution of earnings at event time $t=-3$, that is for women who are three years before giving birth to their first child, in treated vs control municipalities, around the event of a nursery care expansion. In panel (b), we do the same for women at event time $t=-2$. Etc.

Figure C.IV: Pre-School Care



Notes: The figure plots difference-in-differences estimates from specification 6 for the effects of pre-school care expansions on the full dynamics of female earnings. Each panel corresponds to different event time with respect to the birth of the first child. In panel (a) we plot the evolution of earnings at event time $t=-3$, that is for women who are three years before giving birth to their first child, in treated vs control municipalities, around the event of a pre-school care expansion. In panel (b), we do the same for women at at event time $t=-2$. Etc.

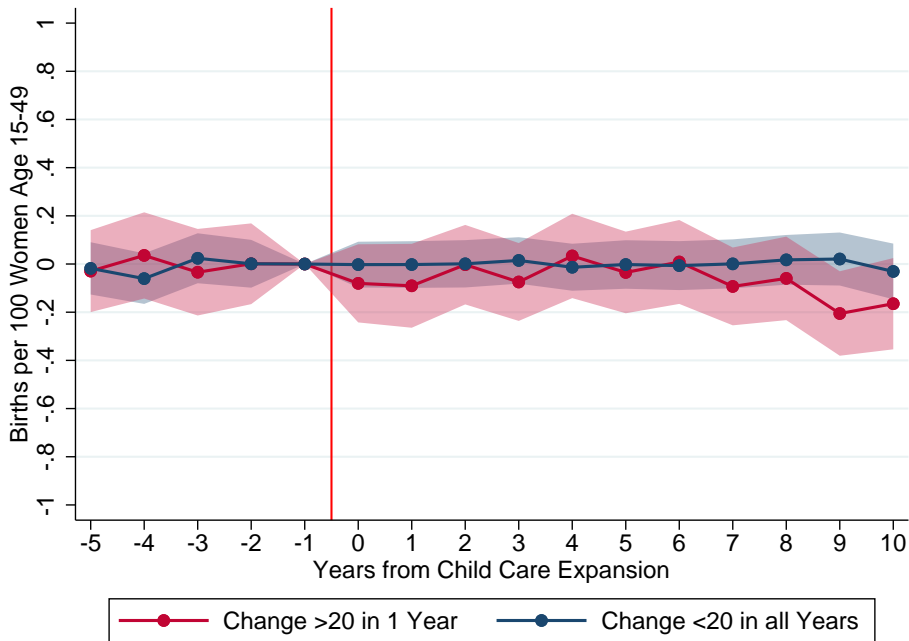
Figure C.V: Effects of 1990 Parental Leave Reform by Level of Child Care Provision



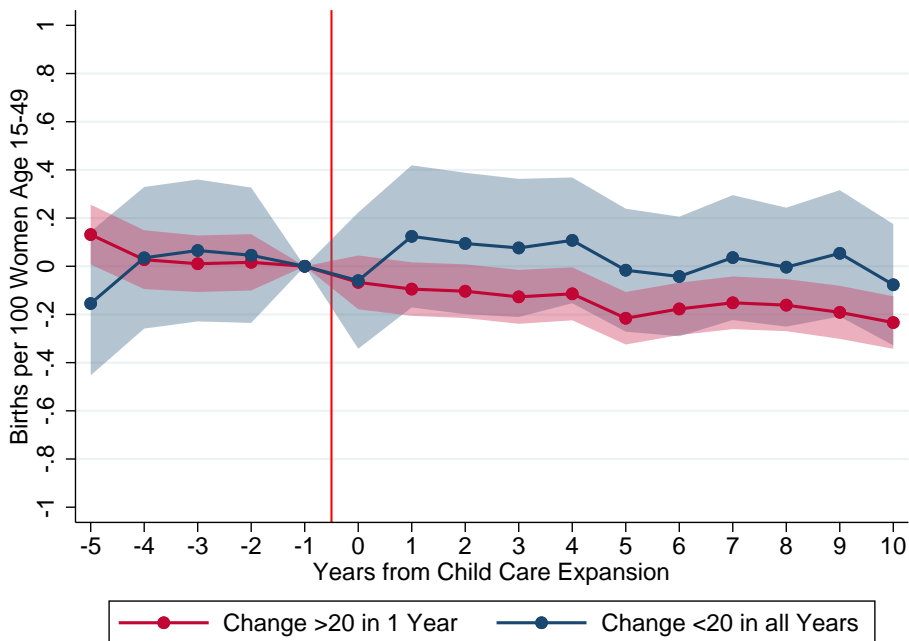
Notes: The figure plots our reform impact estimates by whether the municipality of residence at the time of birth was below or above the median child care index at the time of birth. The reform impact estimates use IPW to weight mothers according to their pre-birth characteristics. Specifically, we run a probit regression of the treatment variable (above median index) on logs of annual earnings in years 3-5 before birth (interacted with year of first birth), employment status in the year before birth and age dummies. We then construct adjusted inverse probability weights based on predicted treatment probabilities and perform the same regressions as in Figure 3.

Figure C.VI: Fertility Effects of Local Child Care Expansions

(a) Nursery Care Expansions



(b) Pre-School Care Expansions



Notes: The figure plots event study estimates of a municipality-level regression of the fertility rate on event time relative to child care expansion interacted with treatment dummy and calendar year dummies. The fertility rate is measured as the number of births in a municipality from the birth register divided by the number of women age 15-49 residing in that municipality in the same year. We obtained the number of women age 15-49 by municipality and year from Statistik Austria register data. Shaded areas are 95% confidence intervals.

D Decomposing Gender Gaps: Data & Methodology

D.1 Measuring Gender Gaps in Earnings Over the Long Run

D.1.1 Using ASSD Data: 1972-2017

Our main data come from the Austrian Social Security Database. This is a matched employer-employee data set which contains detailed information on employment spells and earnings by calendar year over the period 1972 to 2017. The ASSD register covers the universe of paid employment spells, with the notable exclusion of civil servants, self-employed and farmers. This means that about 85% of the Austrian population is covered by ASSD, and this coverage rate has been stable over time. Our analysis therefore focuses on the long run evolution of the earnings gap on the ASSD population.

43

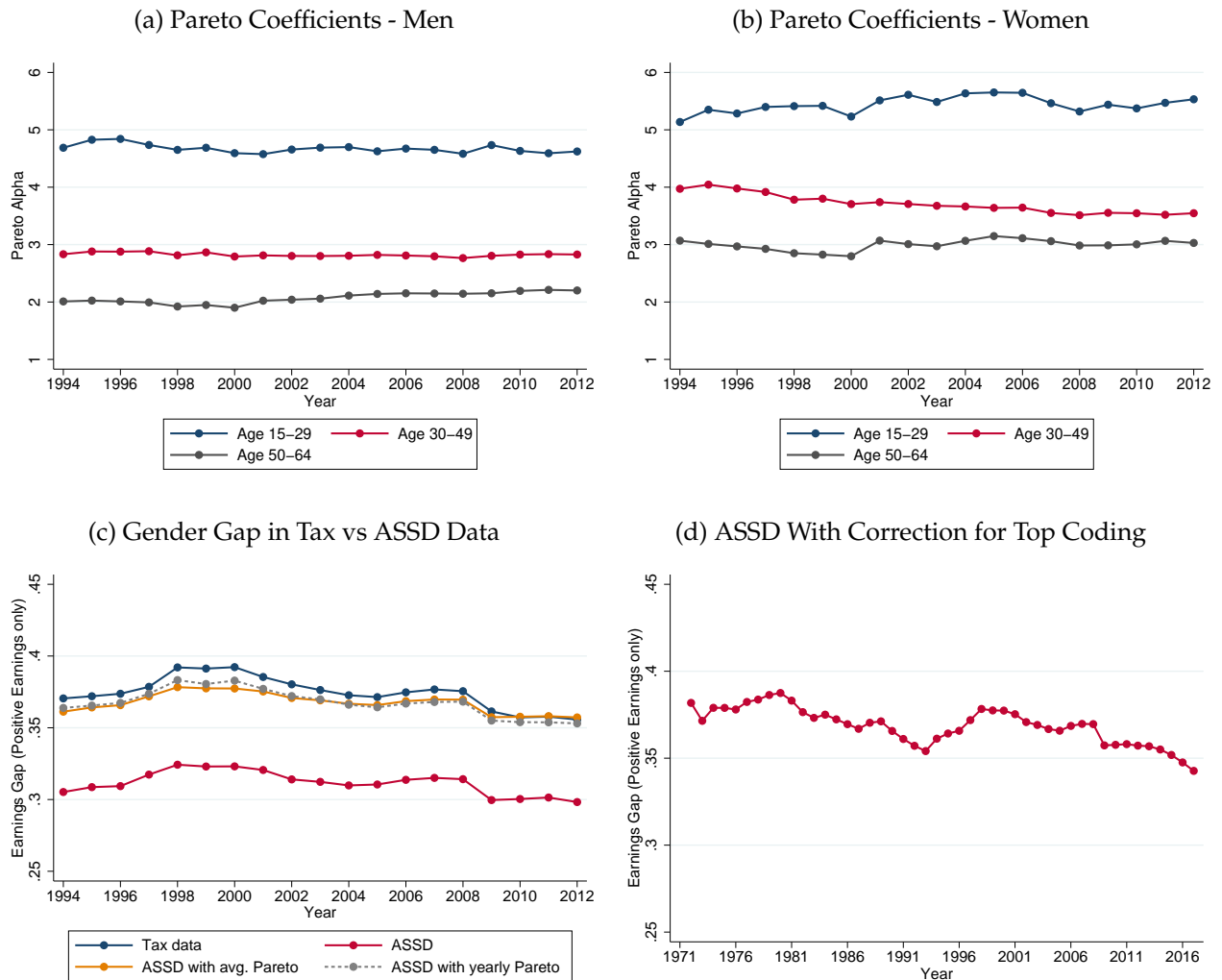
Earnings in the ASSD (and REV) are topcoded at the maximum contribution base, which changes from year to year. The fraction of topcoded individuals is quite stable over the period 1972-2017: around 8% of men and 2% of women have topcoded earnings. During the years 1994-2012 we can check how much topcoding affects earnings gaps by comparing the earnings gaps in the ASSD and in the income tax data (for workers in both datasets). In Figure D.I, panel (c) we plot the earnings gaps calculated from both sources for workers age 15-64, restricting both samples to individual with positive earnings only. Topcoding has mainly a level effect. The ASSD gap is about 12 percentage points below the income tax gap.

To adjust for topcoding in the ASSD data, we impute earnings above the top coding thresholds, using Pareto estimates of the earnings distribution computed using the income tax data. To be precise, we fit, in the tax data, separate Pareto distributions of earnings by gender and age groups. For each age group a and gender g , we compute Pareto coefficients α_a^g . In practice, we use three age groups: 15-29 years old, 30-49 years old, and 50-64 years old. Figure D.I panel (a) reports the Pareto coefficient of these distributions for each year for men, while panel (b) does the same for women. We see that these Pareto coefficients are very stable over time. In panel (c), we impute earnings above the threshold using these estimated Pareto parameters and compute gender gaps adjusted for top coding. The grey dots report estimated gender gaps when we use the yearly Pareto coefficients for each group, while the yellow series shows the gender gaps when we use the average Pareto coefficients over the period 1994-2012. We see that the ASSD series adjusted for top coding match very closely the gender gaps estimated in

⁴³We note that our estimated gender gaps likely understates the earnings gap in the entire Austrian population because men are more often working in the non-ASSD sectors (self employed, civil servants,...) than women and workers in those sectors earn slightly higher wages on average.

the non-top coded tax data, and that using yearly or average Pareto coefficients does not make a significant difference.

Figure D.I: Correcting for Top Coding of Earnings in the ASSD Data



Notes: Panels (a) and (b) plot our estimated Pareto coefficients for men and women by age. This is based on our sample of income tax records (1994-2012) which are not topcoded. We fit a pareto distribution by gender and age groups for gross labour income above the social security limit. The limit is adjusted each year and our procedure only fits the Pareto distribution for income above the relevant cut off in any given year. Panel (c) plots gender earnings gaps calculated from the income tax data, ASSD, ASSD adjusting topcoded earnings based on the average Pareto parameter and ASSD adjusting topcoded earnings based on the relevant Pareto parameter by year. Panel (d) plots the ASSD gender gap in annual earnings (excluding zeroes) based on the ASSD earnings adjusted for topcoding using the 1994-2012 average Pareto parameter.

Our sample includes any individual who has at least one record in the ASSD at any point in time. One issue is that we may miss some individuals who never enter the labor force. In practice, we can compare the fraction of individuals with zero earnings in our

sample and in the census data for years when the census is available. This comparison shows that our sample slightly underrepresents women with zero earnings compared to the census. We therefore added women with zero earnings to our sample to match the share of non-employed in the census.

D.1.2 Measuring Gender Gaps Before 1972

Before 1972, ASSD data exist but has not been digitized. However earnings statistics (*Lohnstufenstatistik*), including detailed tabulations by gender, drawn from the original data, were regularly published, enabling us to consistently compute the evolution of gender inequality on the population covered by the ASSD since 1953.⁴⁴ We used tabulations of the earnings distributions for women and men in 1953, 1957, 1964, 1967, 1970, 1973, 1976, and 1979 that report average monthly earnings for women and men (in July), deciles of the earnings distribution, and the share of women among all ASSD workers. Importantly, to calculate the mean and deciles of all distributions, the IHS imputed topcoded earnings based on the assumption of a lognormal distribution. To account for individuals not in the labor force, we also added men and women with zero earnings so that the employment rates of men and women before 1972 matches the census.

D.2 Decomposition of Gender Gap Between Child Related and Residual Inequality

D.2.1 Methodology

Our methodology for decomposing the gender gap follows [Kleven, Landais and Sørensen \(2019\)](#). It relies on estimating the effect of children (child penalties) on women and men separately. In the spirit of a Oaxaca-Blinder decomposition, we then ask how much of the average gender gap can differences in these estimates explain. We refer to [Kleven, Landais and Sørensen \(2019\)](#) for further details.

We define the average gender gap in year s as $\Delta_s \equiv 1 - E[Y_{ist}^w | s] / E[Y_{ist}^m | s]$. Based on our empirical evidence, and to simplify notation and exposition, we assume in what follows that there are no effect of children on the earnings dynamics of men. We are interested in the average counterfactual gender gap in year s absent child penalties for women, that is: $\tilde{\Delta}_s \equiv 1 - E[\tilde{Y}_{ist}^w | s] / E[Y_{ist}^m | s]$. This gap captures the level of gender inequality that would happen absent child penalties on women's careers. In other words, $\tilde{\Delta}_s$ represents the residual inequality across men and women in the labor market that is not related to

⁴⁴IHS Report ("*Geschlechtsspezifische Einkommensunterschiede: Österreich 1953-1979*") by Christl and Wagner (IHS Vienna 1981)

the arrival of children. The child related gender gap is therefore captured by $\Delta_s - \tilde{\Delta}_s \equiv E[\tilde{Y}_{ist}^w - Y_{ist}^w | s, t] / E[Y_{ist}^m | s]$

In practice, we reconstruct the counterfactual earnings gap in year s $\tilde{\Delta}_s$ based on four inputs:

1. the observed average level of earnings of women in year s , $E[Y_{ist}^w | s]$
2. the fraction ψ_{st} of women who are in event time t with respect to the birth of their first child birth,
3. the child penalty for each event time P_{st} ,
4. and the average level of counterfactual earnings at each event time $E[\tilde{Y}_{ist}^w | s, t]$.

To understand the relationship between the counterfactual earnings gap and these four inputs, note that the computation of $\tilde{\Delta}_s$ solely relies on having an estimate of $E[\tilde{Y}_{ist}^w | s]$. We now show that we can estimate $E[\tilde{Y}_{ist}^w | s]$ based on the four inputs mentioned above. For this, we start by noting that the average level of earnings of women in year s $E[Y_{ist}^w | s]$ is the weighted average of the earnings of women who will never have children (childless women) and of the earnings of women who have, or will have, children and are, in year s , at event time t with respect to the birth of their first child.

$$E[Y_{ist}^w | s] = \psi_s^{Childless} \cdot E[Y_{ist}^w | s, childless] + \sum_t \psi_{st} \cdot E[Y_{ist}^w | s, t]$$

where we have that $\psi_s^{Childless} + \sum_t \psi_{st} = 1$. Similarly, we can write the average counterfactual level of earnings of women in year s $E[\tilde{Y}_{ist}^w | s]$ as the weighted average of the earnings of women who will never have children (childless women) and of the counterfactual earnings \tilde{Y}_{ist}^w of women who have, or will have, children and are, in year s , at event time t with respect to the birth of their first child.

$$E[\tilde{Y}_{ist}^w | s] = \psi_s^{Childless} \cdot E[\tilde{Y}_{ist}^w | s, childless] + \sum_t \psi_{st} \cdot E[\tilde{Y}_{ist}^w | s, t]$$

From this, and the definition of child penalties in equation (2), it follows that:

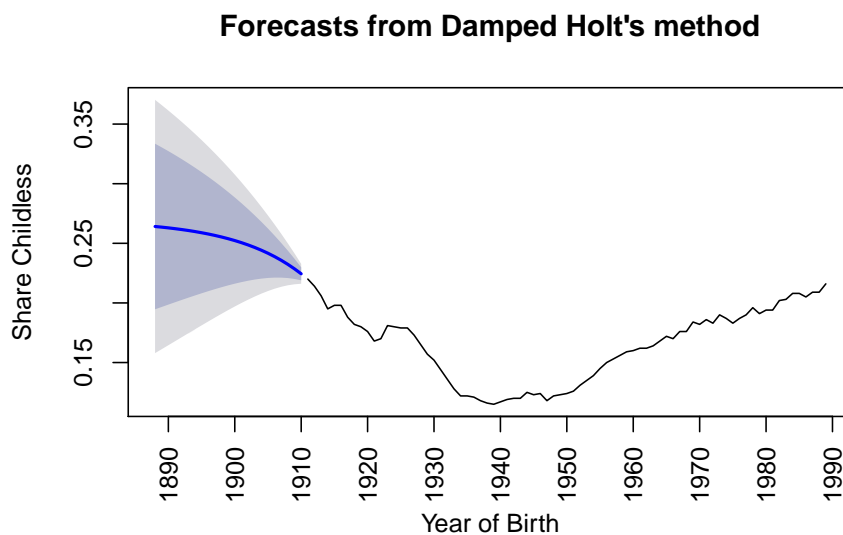
$$\begin{aligned} E[\tilde{Y}_{ist}^w | s] &= E[Y_{ist}^w | s] + \sum_t \psi_{st} \cdot E[\tilde{Y}_{ist}^w - Y_{ist}^w | s, t] \\ &= E[Y_{ist}^w | s] - \sum_t \psi_{st} \cdot E[P_{st} \cdot \tilde{Y}_{ist}^w | s, t] \end{aligned}$$

We show in the following sections how we computed or estimated each of these four inputs.

D.2.2 Estimating Fraction ψ_{st} of Women at Each Event Time

Fraction of Childless Women We estimated the cohort-specific share of childless women from the Human Fertility Database and the Austrian Birth Barometer. Both data sets contain data and estimates by the Vienna Institute of Demography. For older cohorts, we used a simple linear exponential smoothing predictor (Holt's method) to predict childlessness back to cohorts born in 1887 (age 64 in 1951). Figure D.II below shows the evolution of the fraction of childless women, i.e. women whose completed fertility is zero, by birth cohort.

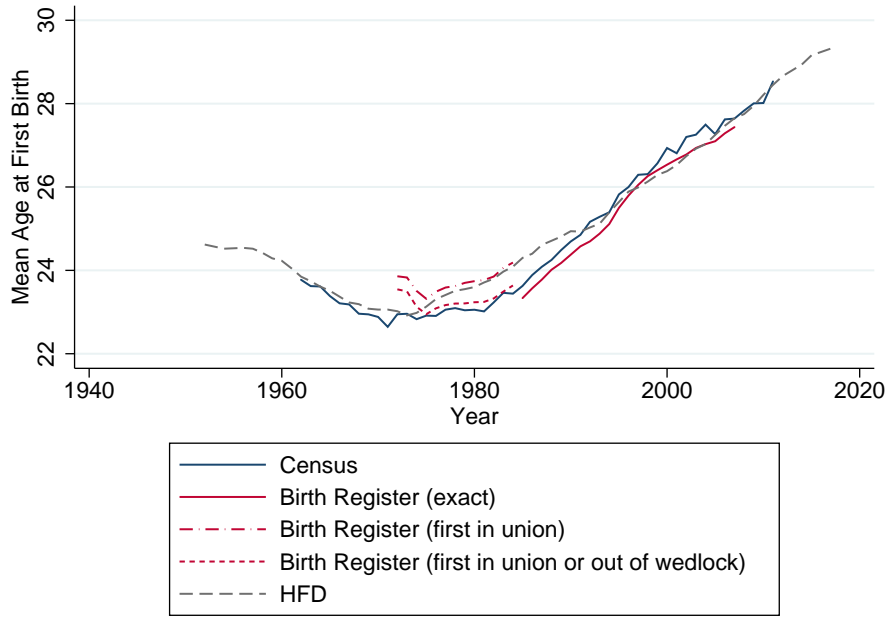
Figure D.II: Evolution of the Fraction of Childless Women by Birth Cohort



Notes: The figure plots the share of childless women (defined as women who will never have children over their lifetime) by birth cohort from the Austrian Birth Barometer (Vienna Institute of Demography) and a backward prediction based on the damped Holt's method ('holt' in R with damped=TRUE and bandwidth 14).

Fraction of Women at Each Event Time To estimate how many mothers are at event time t in a given year s , we need to know the distribution of age at first birth by cohort. Mean age at first birth is available yearly from 1952 to 2017 from the Austrian Birth Barometer (HFD). Figure D.III compares mean age at first birth from the Austrian Birth Barometer and from the Birth Register. We also show age at first birth estimates from IPUMS census microdata (1971-2011).

Figure D.III: Mean Age at First Birth



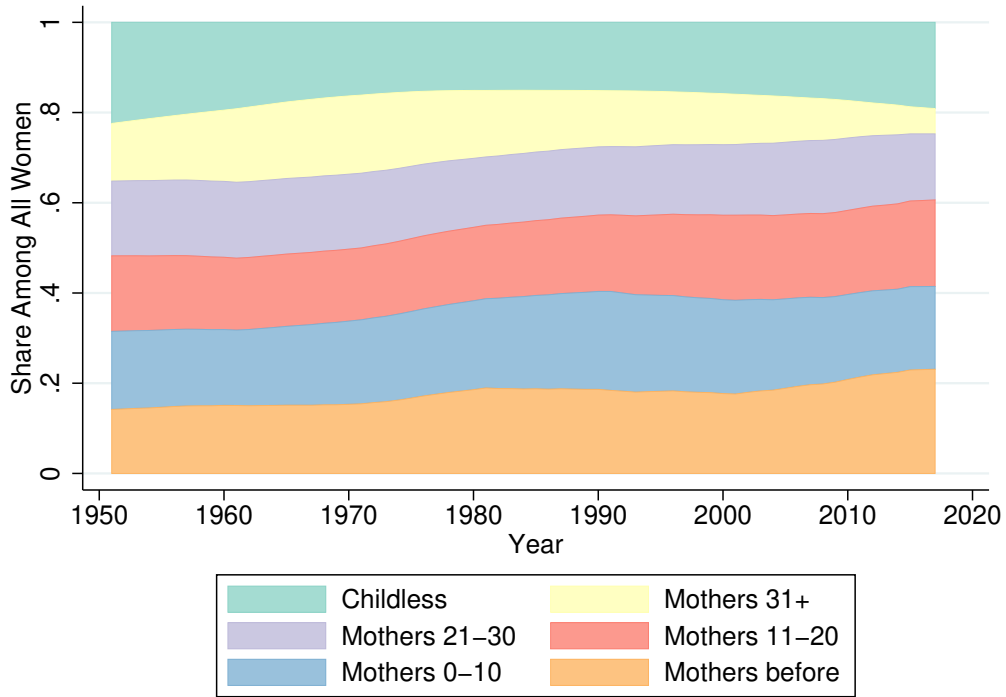
Notes: The figure plots mean age at first birth from various data sources. Census: From the 1971-2011 decennial census (IPUMS and Statistics Austria). We keep all women whose oldest child is between 0-10 at the census date and calculate the average age of the mother in the year the oldest child was born. Birth register (exact): from 1984 the birth register contains data on parity. We keep first births in each year and calculate average age of the mother. Birth register (first in union): before 1984 parity information was only available by mother-father pairs. We keep first births by union in each year and calculate average age of the mother. Birth register (first in union or out of wedlock): We keep first birth by union or births out of wedlock (no father) each year and calculate average age of the mother. HFD: Estimates from the Human Fertility Database.

To get the fraction of women at each event time, we need the distribution of age at first birth, and not just the mean. To this purpose, we fit a Beta distribution to the Birth Register data (1984-2007) for each year. Then we separately regress the two coefficients α_s and β_s of the distribution on mean age at first birth \bar{A}_s . This gives us: $\hat{\alpha}_s = 6.404 - 0.122 \cdot \bar{A}_s$ and $\hat{\beta}_s = 30.346 - 0.966 \cdot \bar{A}_s$.

To predict out of sample we simply plug in mean age at first birth from the HFD and predict α_s and β_s for all years in HFD (1952-2017). We then use the implied Beta distribution in each year to assign mothers to the different event times.

The distribution of age at first birth by year and the childlessness estimates finally enable us to estimate, for each year s the composition of women ($\psi_s^{Childless}$ and $\psi_{st}, \forall t$). The following figure shows the evolution of the composition of women, splitting event times into 5 bins : $t < 0, t = 0 - 10, t = 11 - 20, t = 21 - 30$, and $t > 30$.

Figure D.IV: Composition of Women



Notes: The figure plots our constructed shares of childless women and mothers by event time relative to first birth in each calendar year. See text for details on construction of these shares.

D.2.3 Estimating Child Penalties P_{st} & Counterfactual Earnings Y_{ist}^w By Birth Cohort

We finally need to estimate child penalties and counterfactual earnings by birth cohort over the long run.

To do this, we extend the baseline specification (1) to allow for year-specific coefficients on event time. Specifically we consider the following specification:

$$Y_{ist}^g = \sum_y \sum_{j \neq -1} \alpha_{yj}^g \cdot \mathbf{I}[j = t] \cdot \mathbf{I}[y = s] + \sum_k \beta_k^g X_{kis}^g + v_{ist}^g \quad (10)$$

where we interact the event time dummies with year dummies in order to estimate year-specific event coefficients α_{yj}^g . Note that estimating event coefficients by calendar year s and event year t amounts to estimating event coefficients by birth cohort $c = s - t$. As in our baseline specification (1), we include a full set of age dummies and year dummies in the set of covariates.

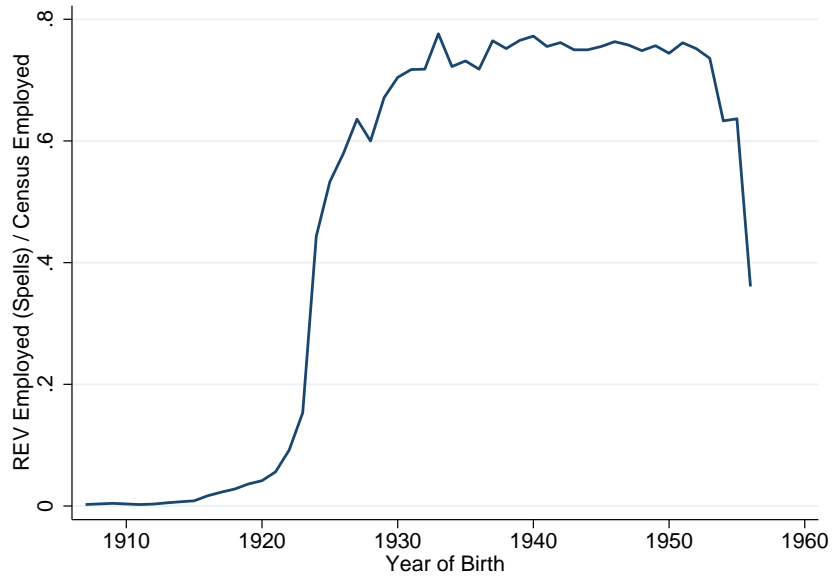
Child Penalties: ASSD Data (1972-2017) and REV data (1960-1972) To compute child penalties by birth cohort since 1960, according to specification (??) we use data from the

ASSD and from the REV before 1972. As explained in section 2, the REV data has a very high coverage rate for women born between the mid-1920s and the 1950s. Panel (a) of Figure D.V below shows that for these birth cohorts, the fraction of women with an employment record represents close to 80% of women with employment records in the census. This number is very similar to the coverage rate of ASSD. But due to the way the REV was created, not all women had their earnings recorded even if the employment and maternity insurance spells are recorded in the REV data. The selection of spells that are recorded with earnings is not random: it is related to the fact that the REV data was set up for pension calculation purposes. When the earnings attached to an employment spell were not relevant for pension computation, i.e. spells with low or no earnings levels (not relevant for top 18 years of earnings variable) the corresponding earnings were less likely to be recorded. This means that spells at younger age with lower earnings level are less likely to have recorded level of earnings attached. As a consequence, the number of first births identified with earnings data in the REV data is lower than the true number of first births, as shown in panel (b) of Figure D.V.

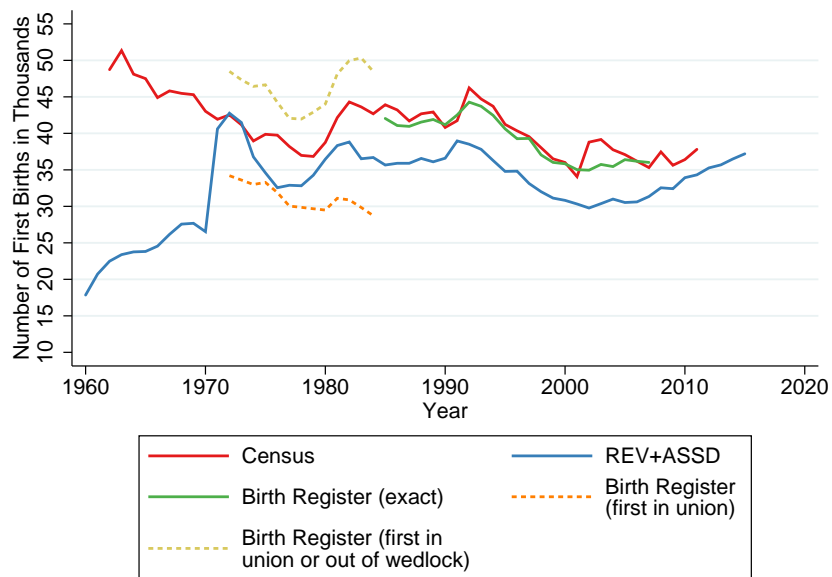
To account for this differential selection of observable first births in the REV data, we use an Inverse Probability Weighting correction. We know that first births are more likely to be recorded with earnings if earnings are relatively high. We therefore reweight the earnings of women who have a first birth in the REV data so that the average earnings of women giving birth relative to average female earnings in Austria in that year matches that observed in the ASSD data. Figure D.VI below shows the evolution of earnings for women observed prior to having children in the REV data up to 1977, and in the ASSD data afterwards. Panel (a) shows the raw data: we see a clear downbreak in the series in 1977, showing that the women who have a first birth spell recorded with earnings information are on average positively selected in the REV data. Panel (b) shows the results when applying our IPW correction.

Figure D.V: Coverage of REV Data

(a) Fraction of Women With Employment Records in REV

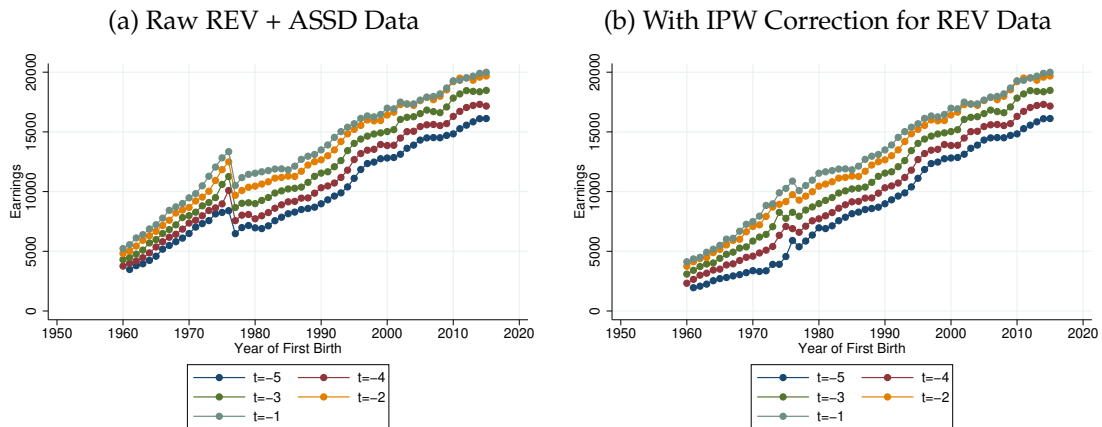


(b) Number of First Births by Year: Comparison Across Data Sets



Notes: Panel (a) plots the share of women with employment records in the REV by year of birth. We select all employment spells from the REV around the 1971 census date (December 5th 1971) and calculate the number of REV employed women by year of birth. We then divide this number by the number of women who are employed according to the 1971 census (again by year of birth). Panel (b) plots the raw number of first birth from various data sources. Census: We keep women with oldest child age 0-10 from each census 1971-2011 and calculate the total number of first births by year. REV+ASSD: We merge the REV and ASSD data and keep the first birth ever observed for each woman. In the REV data this is based on observing a maternity insurance spell, in the ASSD data a live birth is indicated separately. Birth register (various): See notes to Figure D.III for definition of exact and first in union.

Figure D.VI: Earnings by Year of Birth and Event Time in REV and ASSD Data, With and Without IPW Correction



Notes: Panel (a) plots average annual earnings from the REV and ASSD by year of first birth at various event times. We select all first births over the REV+ASSD period. We exclude mothers who never have positive earnings in the REV if their birth was before 1977 to ensure we have 5 years of pre-birth data for everyone in the sample. Panel (b) plots average earnings weighted by IPW. To construct the weights we run yearly probit regressions of 'after' (first birth in 1980 vs e.g. 1970) on the logs of earnings divided by average earnings in the same year from our decomposition procedure at event times -4 to -1 (we include zeroes by adding 1 to earnings) and whether earnings are zero at any of these event times. Based on the predicted probabilities we then construct inverse probability weights.

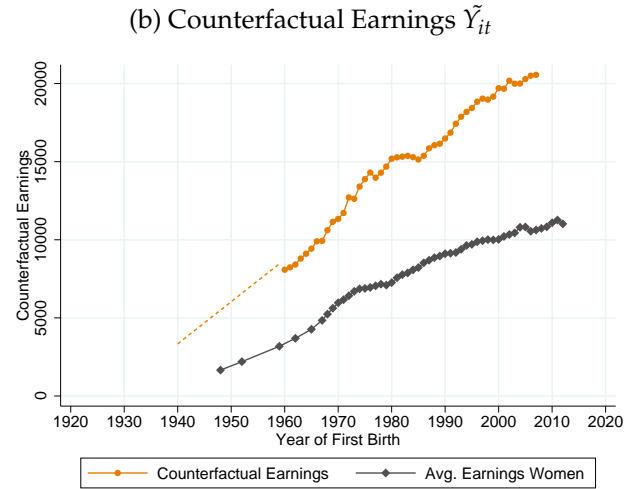
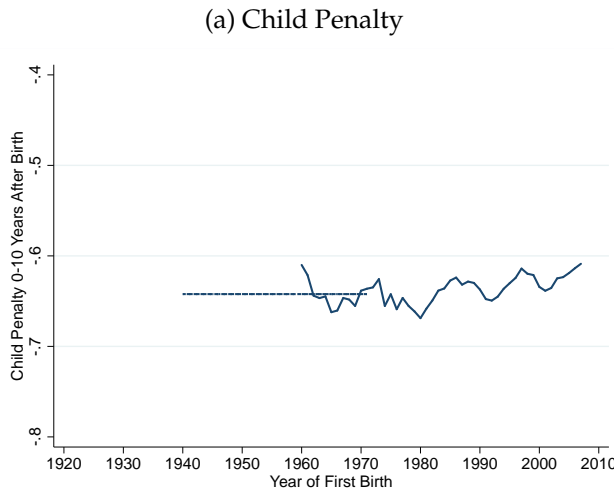
Estimated Child Penalties By Birth Cohorts Figure D.VII shows the estimated child penalties by birth cohort, following specification ??, on the ASSD sample and the REV sample with IPW correction. We also plot the estimated counterfactual earnings $E[\tilde{Y}_{ist}^w | s, t]$ corresponding to these specifications. We see that child penalties in the first ten years after birth have remained remarkably stable over the last 50 years. But child penalties at longer horizons (11 to 20 years after birth, and 21 to 30 years after birth), after being quite stable in the 1960s and early 1970s, have started to decrease significantly.

For first births before 1960, we need to extrapolate child penalties: we assume that child penalties for first births before 1960 were equal to the average penalty observed for first births between 1960 and 1972. For counterfactual earnings corresponding to births before 1960, we extrapolate them linearly based on the observed trend in counterfactual earnings for first births between 1960 and 1972.

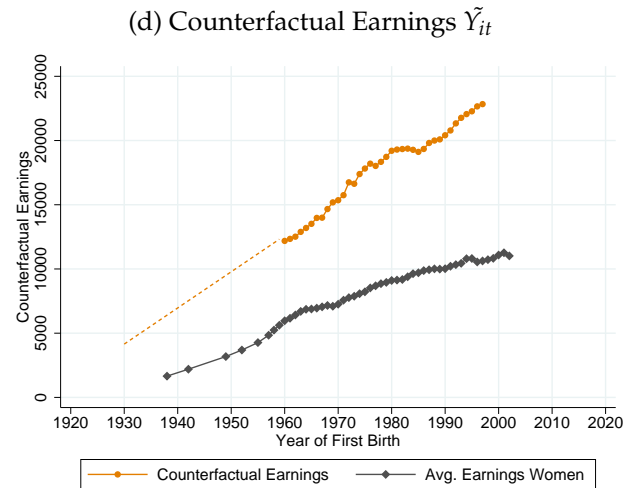
We estimate specification ?? using event time dummies up to $t = 30$. As a result, we also need to extrapolate penalties after event time 30: given the extreme stability of penalties after event time 10, we fix them at the estimated child penalty level in event time 30: P_{s30} .

Figure D.VII: Reforms: Dynamic RD Estimates & Effects on Child Penalties

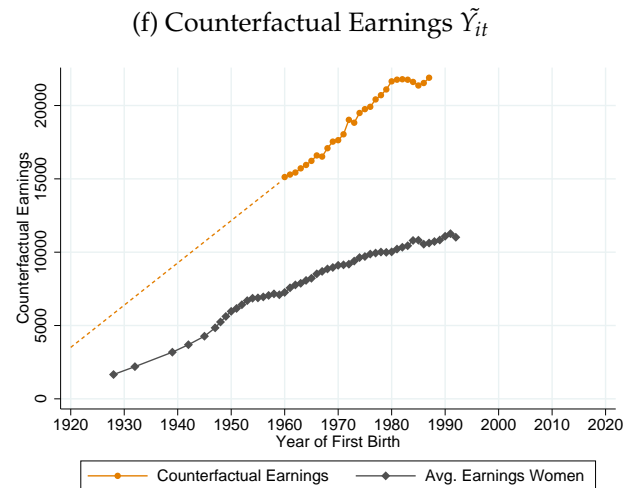
0 to 10 years After Birth



11 to 20 years After Birth



21 to 30 years After Birth



Notes: The figure plots average child penalties and predicted counterfactual earnings over event times 0-10, 11-20 and 21-30 by year of first birth (see text for details). Dashed lines indicate the extrapolations we use based on simple averages for child penalties (1960-72 average) and linear backward prediction for counterfactual earnings. We also add our constructed average earnings of all women (including zeroes) in the figures on the right.

D.3 Historical Gender Gaps With Counterfactual Family Policies

We now turn to explaining how we can use the previous decomposition, and our estimates of the effects of policies on child penalties to measure gender gaps under counterfactual policy scenarii.

Our focus is on the three following family policies: parental leave duration (τ_{PL}), the coverage of free nursery care (τ_N), and the coverage of free pre-school care (τ_{PS}). What we are trying to identify, for family policy $\tau_k \in \{\tau_{PL}, \tau_N, \tau_{PS}\}$, is the causal effect of any policy change $d\tau_k$ on the level of the gender gap: $\frac{d\Delta_s}{d\tau_k}$. Based on our previous decomposition exercise, we can express this effect as a simple function of the causal effects of the policy change $d\tau_k$ on the full dynamics of the child penalty ($dP_t/d\tau_k, \forall t$). Indeed, we have that:

$$\begin{aligned} \frac{d\Delta_s}{d\tau_k} &= -\frac{dE[Y_{ist}^w|s]/d\tau_k}{E[Y_{ist}^m|s]} \\ &= -\frac{1}{E[Y_{ist}^m|s]} \cdot \left\{ \sum_t \psi_{st} \cdot \frac{dP_t}{d\tau_k}(\tau_{PL}, \tau_N, \tau_{PS}) \cdot E[\tilde{Y}_{ist}^w|s, t] \right\} \end{aligned}$$

To get the above expression, we used the empirical fact that earnings of men are affected neither by the arrival of kids, nor by family policies. The above expression also assumes that family policies have no effect on fertility, such that $d\psi_{st}/d\tau_k = 0$. But we explore below scenarii where we relax this assumption and allow family policies to affect fertility as well.

Note that the effect of any policy change on the child penalty is theoretically a function of *all* family policies. But in practice, our empirical evidence found no interaction effects between parental leave and child care provision. This means that we can consider the impact of a policy change on the child penalty to be solely a function of that particular policy: $\frac{dP_t}{d\tau_k}(\tau_k)$.

As a consequence, for all family policy reforms since 1961, we can compute the causal effect of that reform on the gender gap in each year s by simply combining:

1. our estimates of ψ_{st} , the fraction of women at each event time t in year s
2. our estimates of average counterfactual earnings $E[\tilde{Y}_{ist}^w|s, t]$ for women absent child penalties, at each event time t in year s
3. and our estimates of $\frac{dP_t}{d\tau_k}(\tau_k)$, the effect of each the reform on child penalty at each event time t

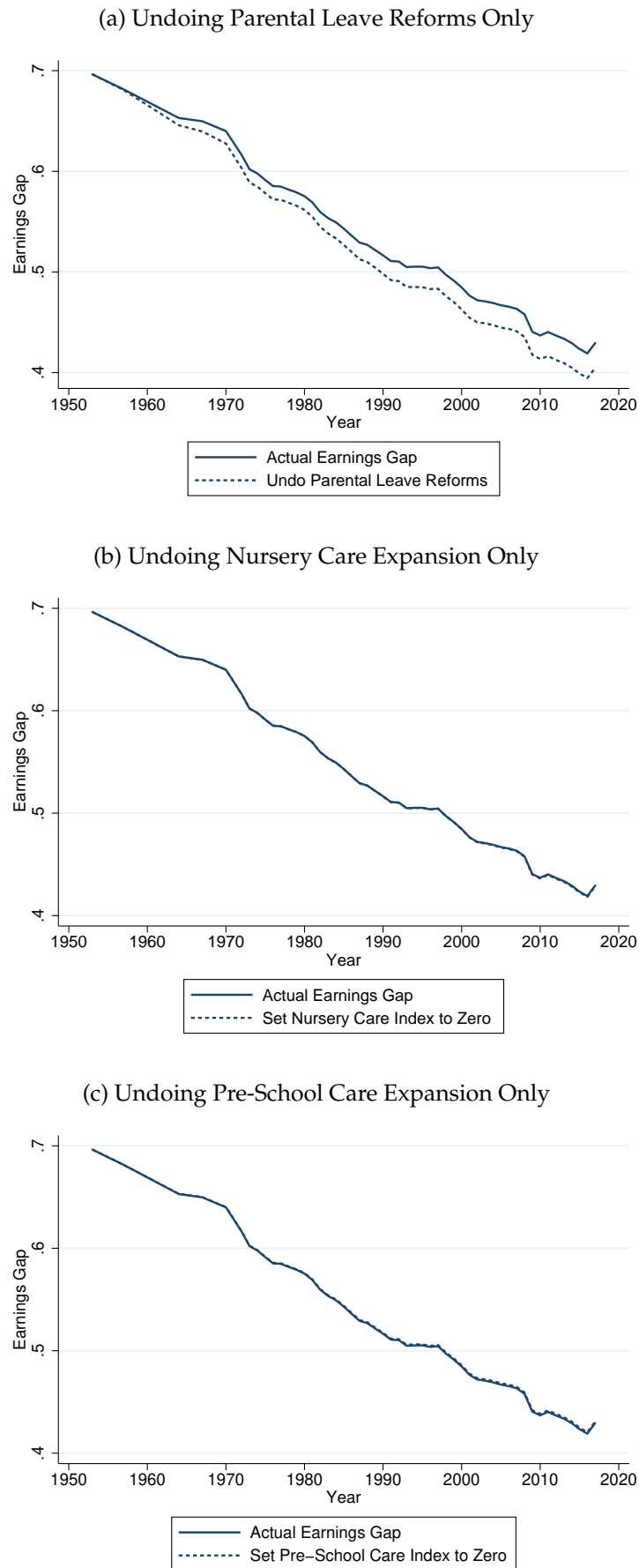
Baseline Results: Undoing All Family Policies Our baseline counterfactual policy scenario consists in getting rid of all family policies. In other words, we simulate the evolution of the gender gap since 1961, when eliminating the effects of all parental leave reforms, and setting the level of nursery care and of pre-school care provision to zero. The table below summarizes our estimates of the effect of each policy on child penalties that we use to compute the counterfactual gender gaps. Note that for parental leave reforms, we account for the steady state effect of the reform on child penalties, that is we account for the effect on child penalties of getting access to a new regime PL regime for all children, and not just the first one. This issue is discussed at length in section 3.3 and results of these steady state effects on child penalties are presented in Figure B.IX.

	Avg. effect on 0-10 Penalty (first birth)	Avg. effect on 0-10 Penalty (higher order births)
1961 Reform	-0.0281	-0.0459
1990 Reform	-0.0227	-0.0234
1996 Reform	0.0109	0.0251
2000 Reform	-0.0149	-0.0145
Index 1-2 Expansion (for Index from 0 to 100)		0.022
Index 3-5 Expansion (for Index from 0 to 100)		0.027

Note also that given we systematically found no long run effects of any of these policies on child penalties, we systematically set the effect of each policy reform on penalties to be zero for event time superior to 10: $\frac{dP_t}{d\tau_k} = 0, \forall t > 10$.

Figure D.VIII below shows the effect on the gender gap of removing each policy separately. Panel (a) shows what the gender gap would have been in the absence of any parental leave reform. In this scenario, we only undo parental leave expansion, and assume that child care provision would have followed the same path as actually observed. The graph highlights that parental leave expansions have had a negative, albeit small, impact on gender inequality. This result stems from the significant negative impact of parental leave policies on labor supply in the short run. But the absence of long run effects of PL on the careers of women means that the overall negative impact of parental leave policies on gender inequality remains somewhat modest. While parental leave policies do have a small effect on the evolution of the gender gap, we find in panel (b) and (c), to the contrary, that child care policies have none. When undoing these policies, we find that the gender gap would have been absolutely unaffected by the absence of nursery care or of pre-school care.

Figure D.VIII: Gender Gap in Earnings and Counterfactual Gender Gaps Under Different Policy Scenarii



Notes: The figures plot actual earnings gaps and counterfactual earnings gaps based on setting various policy parameters to zero in isolation. See text and notes to Table 1 for details.

Sensitivity Analysis It is critical to assess how sensitive our counterfactual gender gaps estimates are to the precision with which we estimate the effects of policies on child penalties. To this effect, and to provide some guidance for inference, we computed two bounding counterfactual gender gaps. The first, “optimistic” one, uses, instead of our point estimates, the upper bound of the 95% confidence interval on $\frac{dP_t}{d\tau_k}(\tau_k)$ for each policy change. The second, or “pessimistic” scenario, uses the lower bound of the 95% confidence interval. The results of these two bounding scenarii are presented in Table 1. The bottom-line is that our null effects are relatively precise: we can clearly rule out large positive effects of family policies on gender inequality.

Accounting for Fertility Effects We assumed, for our baseline counterfactual policy scenario that family policies had no effect on fertility. We now explore scenarii allowing for potential effects on fertility. In practice, this means accounting for potential effects of each policy change on the fraction of women observed at each event time t : $\frac{d\psi_{st}}{d\tau_k}$. The causal effect of the policy change on the gender gap can now be written as:

$$\begin{aligned} \frac{d\Delta_s}{d\tau_k} &= -\frac{dE[Y_{ist}^w|s]/d\tau_k}{E[Y_{ist}^m|s]} \\ &= -\frac{1}{E[Y_{ist}^m|s]} \cdot \left\{ \sum_t \psi_{st} \cdot \frac{dP_t}{d\tau_k} \cdot E[\tilde{Y}_{ist}^w|s, t] + \sum_t \frac{d\psi_{st}}{d\tau_k} \cdot P_{st} \cdot E[\tilde{Y}_{ist}^w|s, t] \right\} \end{aligned}$$

where the second part in the bracketed term accounts for fertility responses. Intuitively, if a policy reform induces more couples to have children, it will increase the fraction of women observed at each event time t after birth ($d\psi_{st}/d\tau_k > 0$). As these women are facing penalty P_{st} , this will increase the gender gap.

We use two different sources to account for fertility effects. To account for fertility effects of parental leave we use the estimates from [Olivetti and Petrongolo \(2017\)](#) Table 3 column (8). Their formula for total fertility rate is (from the replication data set)

$$TFR = \dots - 0.0007448 \times \text{leave_weeks} + 0.0013043 \times \frac{\text{leave_weeks}^2}{100} + 0.0019767 \times \% \text{weeks_paid} + 0.0003374 \times \% \text{payrate}$$

The following table shows what these estimates imply when we undo parental leave. Note that in their data set `leave_weeks` is equal to weeks of maternity insurance plus weeks of parental leave. So undoing parental leave but keeping maternity insurance means setting leave duration to 8 weeks.

To account for fertility effects of **child care expansions** we use our own estimates of the effect of these roll-outs on total fertility rate. We constructed total fertility rate at the mu-

Table D.I: Summary of Fertility Impact Computed Based on Estimates from [Olivetti and Petrongolo \(2017\)](#)

Regime	leave (wks)	%wks paid	%payrate	Counterfactual			TFR Effect of PL
				leave (wks)	%wks paid	%payrate	
1961-1990	60	100	52	8	100	100	-0.0088
1990-1996	112	100	46	8	100	100	0.0671
1996-2000	112	75	37	8	100	100	0.0146
2000-2008	112	123	30	8	100	100	0.1072

Notes: The table shows the policy parameters we use for the different periods to calculate the total fertility effect of parental leave policies. We use the formulas in the online dofiles provided by [Olivetti and Petrongolo \(2017\)](#).

municipality level as follows. First, we get the number of births at the municipality level from the birth register (1988-2007) and Statistik Austria Register Data (2002-2018).⁴⁵ Then we get the number of women age 15-49 from the census (1991, 2001, 2011) and the Population Register (2011-2017). We interpolate linearly to get yearly data. Dividing the number of births by the number of women is the crude fertility rate. We then use the Austrian aggregate yearly relationship between crude fertility rate and total fertility rate to convert crude to total fertility rate.

We use the effects 10 years after expansion as an estimate for the TFR impact of child care expansion, scaled by the difference in the index during event times 0-5.

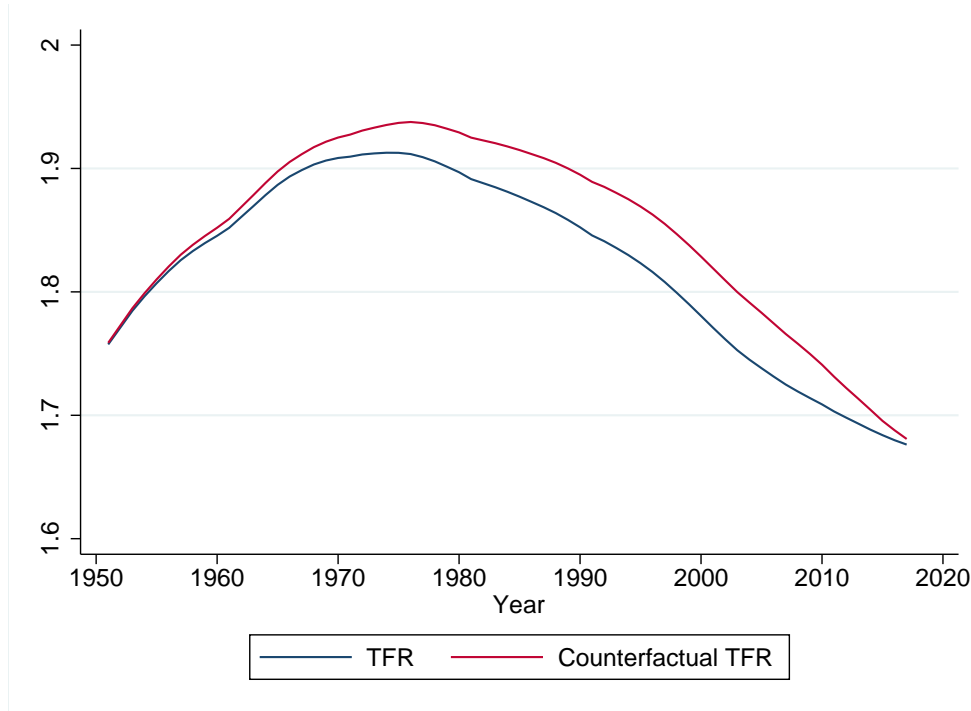
	Effect at $t = 10$	Avg. change in index	Scaled Effect
Nursery Care Expansion	-0.049645	0.363941	-0.1364
Pre-School Care Expansion	-0.046957	0.369499	-0.1271

Figure [D.IX](#) shows actual and counterfactual total fertility rates. Counterfactual means fertility when there is no parental leave and no child care.

⁴⁵And we take the average of those two numbers for years 2002-2007 where the data overlaps, to account for a few small discrepancies in the data.

Figure D.IX: Actual and Counterfactual Total Fertility Rates

(a) By Year



Notes: The figure plots the total fertility rate (TFR) calculated by multiplying the number of children by the shares of women who have 0-4 children as constructed for the main decomposition exercise. The counterfactual TFR is based on setting all family policy parameters to zero taking into account our estimated child care expansion fertility effects and the parental leave expansion effects based on [Olivetti and Petrongolo \(2017\)](#).

To convert our estimated fertility effects into estimates for $\frac{d\psi_{st}}{d\tau_k}$, we proceed as follows. Our fertility estimates are, theoretically, measuring the effect of policy reforms on total fertility $dN/d\tau_k$, where N is the average number of children by women. Note that $N = (1 - \psi^{childless}) \cdot \bar{n}$, where \bar{n} is the average number of children among women with children and $\psi^{childless}$ is the fraction of childless women.

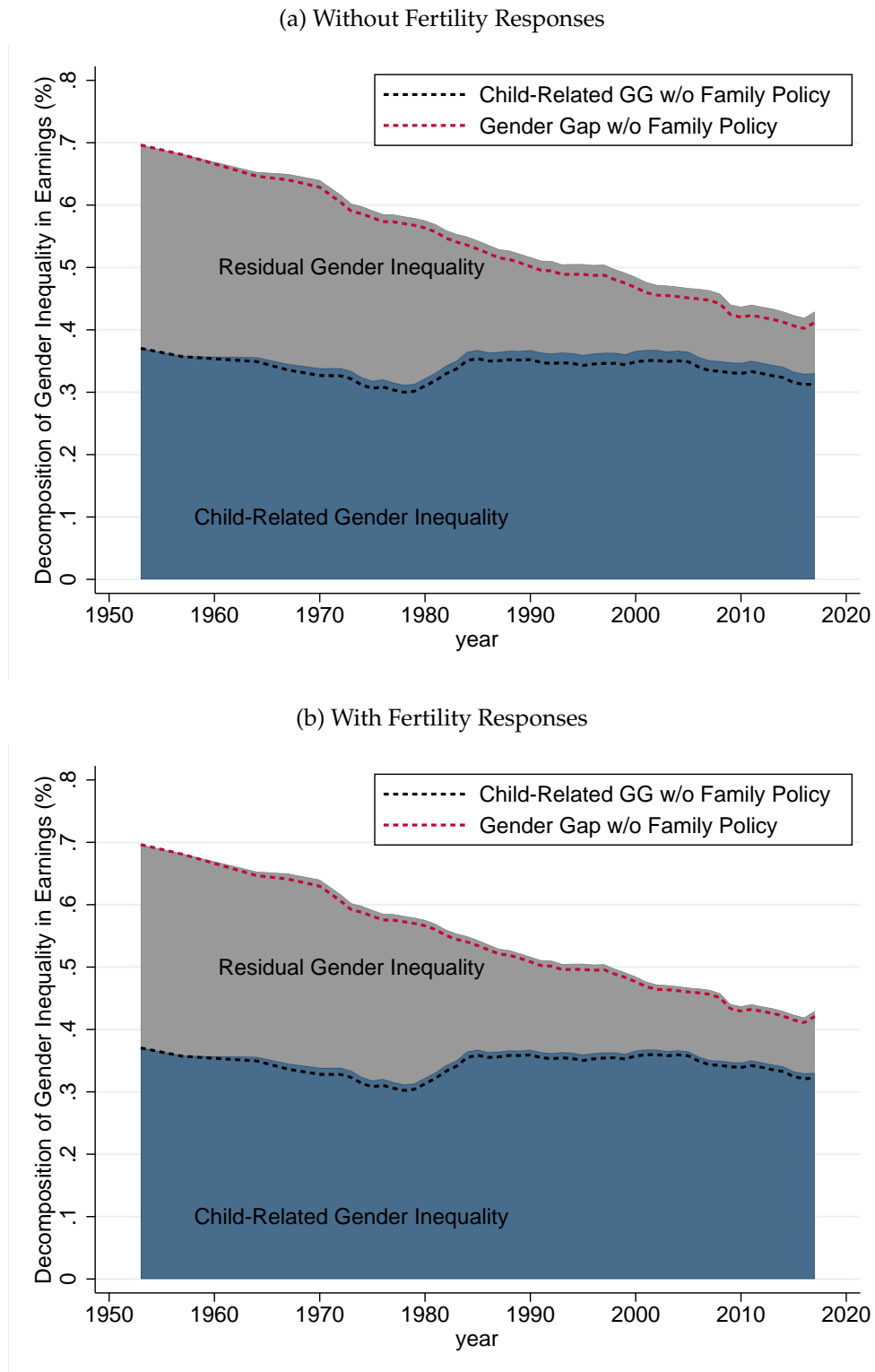
What we do is to assume that policy reforms affect the total fertility by changing the fraction of women with kids rather than the number of kids per mother. In other words, we assume that $dN/d\tau_k = d(1 - \psi^{childless})/d\tau_k \cdot \bar{n}$

As a consequence, we easily obtain an estimate of the effect of the reform on the fraction of women at event time t in year s :

$$\begin{aligned} \frac{d\psi_{st}}{d\tau_k} &= -\frac{d(\psi^{childless})}{d\tau_k} \cdot \psi_{st} \\ &= \frac{dN}{d\tau_k} \cdot \frac{\psi_{st}}{\bar{n}} \end{aligned}$$

Figure D.X shows the results of our counterfactual policy scenario of removing all family policies, when accounting for fertility responses. For almost the entire period the policy reforms slightly decreased fertility overall, so undoing the reforms slightly increases the gender earnings gap. In Table D.III we also explore the sensitivity of our results to the precision with which the fertility effects are estimated. We computed two bounding counterfactual gender gaps. The first, “optimistic” one, uses, instead of our point estimates, the upper bound of the 95% confidence interval on $\frac{dN}{d\tau_k}$ and $\frac{dP_t}{d\tau_k}(\tau_k)$ for each policy change. The second, or “pessimistic” scenario, uses the lower bound of the 95% confidence interval.

Figure D.X: Counterfactual Gender Gaps Under No Family Policy Accounting for Fertility Responses



Notes: The figures plot the results of our decomposition exercise removing all family policies when we do not take fertility effects into account (panel (a)) and when we incorporate fertility effects (panel (b)). See text for details.

Table D.III: Sensitivity Of Counterfactual Gender Gap Estimates

Year	1953	1964	1970	1980	1990	2000	2010
Panel A. Decomposition							
Actual Earnings Gap	0.696	0.653	0.640	0.575	0.517	0.485	0.437
Counterfactual Earnings Gap	0.326	0.297	0.302	0.253	0.149	0.119	0.090
Gap Due to Kids	0.370	0.356	0.338	0.322	0.367	0.366	0.347
Panel B. Undoing the Reforms Without Fertility Response							
<i>No Parental Leave and No Child Care (Baseline)</i>							
Earnings Gap	0.696	0.646	0.628	0.562	0.499	0.463	0.414
Child-Related Gender Gap	0.370	0.349	0.326	0.309	0.349	0.344	0.324
<i>No Parental Leave and No Child Care (Optimistic)</i>							
Actual Earnings Gap	0.696	0.648	0.632	0.565	0.503	0.474	0.428
Child-Related Gender Gaps	0.370	0.351	0.330	0.312	0.353	0.355	0.338
<i>No Parental Leave and No Child Care (Pessimistic)</i>							
Earnings Gap	0.696	0.644	0.624	0.559	0.495	0.452	0.400
Child-Related Gender Gap	0.370	0.346	0.322	0.306	0.345	0.333	0.309
Panel C. Undoing the Reforms With Fertility Responses							
<i>No Parental Leave and No Child Care (Baseline)</i>							
Earnings Gap	0.696	0.646	0.629	0.565	0.505	0.472	0.423
Child-Related Gender Gap	0.370	0.349	0.327	0.312	0.356	0.353	0.332
<i>No Parental Leave and No Child Care (Optimistic)</i>							
Earnings Gap	0.696	0.649	0.633	0.568	0.510	0.483	0.437
Child-Related Gender Gap	0.370	0.352	0.331	0.315	0.360	0.364	0.347
<i>No Parental Leave and No Child Care (Pessimistic)</i>							
Earnings Gap	0.696	0.644	0.625	0.561	0.501	0.460	0.408
Child-Related Gender Gap	0.370	0.347	0.323	0.308	0.352	0.341	0.318

Notes: The table reports inputs and results of our decomposition exercise of gender gaps for selected years, when accounting for fertility responses to family policies. See notes to Table 1 for additional details. The “optimistic” scenario is based on the upper bound of the 95% confidence interval for all policy treatment effect estimates and the “pessimistic” scenario is based on the lower bound of the 95% confidence interval.