American Economic Journal: Economic Policy 2024, 16(2): 110–149
https://doi.org/10.1257/pol.20210346

Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation†

By Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller*

Do family policies reduce gender inequality in the labor market? We contribute to this debate by investigating the joint impact of parental leave and childcare, using administrative data covering 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 to compute counterfactual gender inequality series. Our results show that the enormous expansions of parental leave and childcare have had virtually no impact on gender convergence. (*JEL* D63, J13, J16, J31, J32)

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 Søgaard 2019; Kleven et al. 2019), which suggests that family policies such as parental leave (PL) and childcare have the greatest potential to affect gender inequality. A sizable 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 PL and childcare over more than half a

* Kleven: Princeton University, NBER, CEPR, and CEBI (email: kleven@princeton.edu); Landais: London School of Economics and CEPR (email: c.landais@lse.ac.uk); Posch: Analysis Group (email: johanna.posch@analysisgroup.com); Steinhauer: University of Edinburgh and CEPR (email: andreas.steinhauer@ed.ac.uk); Zweimüller: University of Zurich and CEPR (email: josef.zweimueller@econ.uzh.ch). Matthew Notowidigdo was coeditor for this article. 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 #DNRF134 from the Danish National Research Foundation. Landais gratefully acknowledges support from the LSE International Inequality Institute and from ERC grant #679704-DYNAMICSS.

† Go to https://doi.org/10.1257/pol.20210346 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

century. Our main conclusion is simple and striking: the enormous expansions of PL and childcare 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. PL 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 childcare was rolled out across the country. Similar expansions of family policies have been implemented in other European countries, and their implications for gender inequality are the subject of heated debate: Is access to job-protected PL and subsidized childcare instrumental for boosting female labor force participation and reducing gender gaps? Are generous PL 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 et al. 2019; Boelmann, Raute, and Schönberg 2020)?[1]

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 PL and childcare policies. Family policies are often designed as a bundle, with childcare coverage starting around the time that PL eligibility is ending. As a result, we analyze the effects of PL and childcare policies jointly, capturing any nonlinearities brought about by their potential interaction.

Leveraging a regression discontinuity (RD) design based on cutoff dates of the PL reforms, we first show that PL duration has significant negative effects on the earnings of women relative to men in the short run. We show that these effects are nonlinear, being greater at lower baseline durations. However, for all PL reforms, we find precisely estimated zero effects in the long run. Leveraging a difference-in-differences (DiD) design similar to Duflo (2001) based on large expansions in locally provided childcare, we find precisely estimated zero effects

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

of both nursery care and preschool care on the earnings of women relative to men.[2] We provide suggestive evidence that this null effect is driven by the 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 childcare and PL 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 Søgaard (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 a remarkably small impact on gender inequality. Due to the negative effects of PL on labor supply in the short run, family policies have been slightly counterproductive for gender convergence. The average earnings of Austrian women are 43.5 percent lower than the average earnings of men today, while they would have been 41.5 percent 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 Søgaard 2019; Kleven et al. 2019; Kuziemko et al. 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 PL schemes (see, e.g., Lalive and Zweimüller 2009; Schönberg and Ludsteck 2014; Dahl et al. 2016) and childcare subsidies (see, e.g., Baker, Gruber, and Milligan 2008; Havnes and Mogstad 2011a). 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

[2] Similar designs using local supply shocks have been used to study the impact of childcare on female employment (e.g., Havnes and Mogstad 2011a) and children's outcomes (e.g., Cornelissen et al. 2018).

[3] The absence of any interaction effects between childcare provision and PL is relatively unsurprising ex post (but not ex ante) given our finding that the effects of the two policies separately are zero.

[4] Our null result is striking given that the subsidies to PL and childcare 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 childbirth. 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 Earned Income Tax Credit program in the United States. See Kleven (2014) for a cross-country analysis of participation tax rates and labor supply.

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 I describes the Austrian institutional context and data. Section II presents quasi-experimental evidence on the impact of PL reforms on child penalties. Section III presents quasi-experimental evidence on the impact of local childcare expansions on child penalties. This section also provides evidence on the potential mechanisms behind our null effect. Section IV lays out our decomposition framework and provides estimates of the counterfactual evolution of gender inequality under different policy trajectories. Section V concludes.

I. Context and Data

A. *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 percent and the female labor force participation rate is about 55 percent (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 United States and the United Kingdom have broadly similar levels of gender inequality in earnings. Moreover, female labor force participation in Austria is higher than it is 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. Online Appendix Figure A1 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 on the attitudes toward market work of 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 United States, the United Kingdom, and France. Overall, the evidence in online Appendix Figure A1 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: PL and childcare provision (Olivetti and Petrongolo 2017). Austria is a case in point.

In the aftermath of World War II, there was no system of PL in Austria. Women were entitled to Mutterschutz, a maternity insurance period of eight 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 PL policies (in 1990, 1996, 2000, and 2008), which have considerably expanded the generosity and scope of PL.[5] Today, the Austrian PL 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 PL 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 percent of net earnings. The take-up of PL comes with a job-protection guarantee of 24 months. The receipt of PL benefits is not subject to any work eligibility requirement prior to childbirth.

Institutional childcare in Austria is divided into nursery care (Krippe) covering children from ages 1 to 2 and preschool care (Kindergarten) from ages 3 to 5. The provision of both forms of institutional childcare has been greatly expanded over the past 50 years. In the beginning of the 1970s, fewer than 2 percent of children between the ages of 1 and 2 attended nursery. Today, this number has risen to 25 percent. The fraction of children attending preschool has risen from 40 percent to more than 90 percent over the same time period. Childcare is provided by municipalities, following general guidelines set by state legislation.[6] The median number of daily operating hours for Austrian childcare institutions is nine hours, with typical opening hours from 7:30 AM to 4:30 PM.[7] Institutional childcare 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 percent of family net income in 2015, compared to an average of 13 percent across the OECD and 22 percent in the United States.[8] Note that subsidized institutional childcare encompasses both public- and church-provided care. The private market for childcare is very limited and essentially consists of a few private creches and childminders (Tagesmutter).[9]

[5] Two 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.

[6] State 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).

[7] See 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.

[9] Figures from a microcensus survey in 2002 reveal that less than 3 percent of children under age two were taken care of by Tagesmutter.

B. *Data*

Labor Market Outcomes and 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 et al. 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, the self-employed, and farmers. This means that about 85 percent of the Austrian population is covered by the 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 percent 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 childbirth events of women (e.g., date of birth of each child).

ASSD data exist before 1972 but have 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 and 1972, the REV covers about 80 percent of workers, as shown in online Appendix Figure D.V. Although the REV does not contain direct information on childbirths, 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 enable 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.

Importantly, both the ASSD and REV contain childbirth information only for women. We can nevertheless link children to their fathers for a large subset of childbirths. To this end, we use the child benefit register (Beihilfedaten). This register contains the universe of child benefit claims filed between 1995and 2012. As the take-up of child benefits is almost complete, we are able to match around

[10] We provide further details on our methodology and its robustness in online Appendix Section D.1.

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

90 percent of children born between the 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.

Childcare Provision.—Our analysis also relies on granular information on childcare 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 childcare facilities (broken down by nurseries and preschools), 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 childcare supply at the municipal level. We provide details about the construction of these measures in Section III.

C. *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 Søgaard (2019). For each parent in the data, event time t is indexed relative to the year of the first childbirth. 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:

$$Y_{ist}^{g} = \boldsymbol{\alpha}^{g}\mathbf{D}_{ist}^{Event} + \boldsymbol{\beta}^{g}\mathbf{D}_{ist}^{Age} + \boldsymbol{\gamma}^{g}\mathbf{D}_{ist}^{Year} + \nu_{ist}^{g}. \tag{1}$$

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] 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 life cycle 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]

[12] For 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 $\mathbf{D}$.

[13] Kleven, Landais, and Søgaard (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.

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:

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

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 percent 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 online Appendix Figure A.II, long-run earnings penalties in Scandinavia lie in the 20–30 percent 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.

II. The Impact of Parental Leave on Child Penalties

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

As a starting point, we focus on the PL regime in place since 2008. This regime offers parents the possibility of choosing between three PL options: (i) a long PL ("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 PL ("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 PL ("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

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

[15] Importantly, under the identifying assumptions, equation (2) corresponds to $\hat{\alpha}_t^g / E\left[\tilde{Y}_{ist}^g \middle| t\right] = \left(E\left[y_{1,t} \middle| t\right] - E\left[y_{0,t} \middle| t\right]\right) / E\left[y_{0,t} \middle| t\right]$, 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.

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

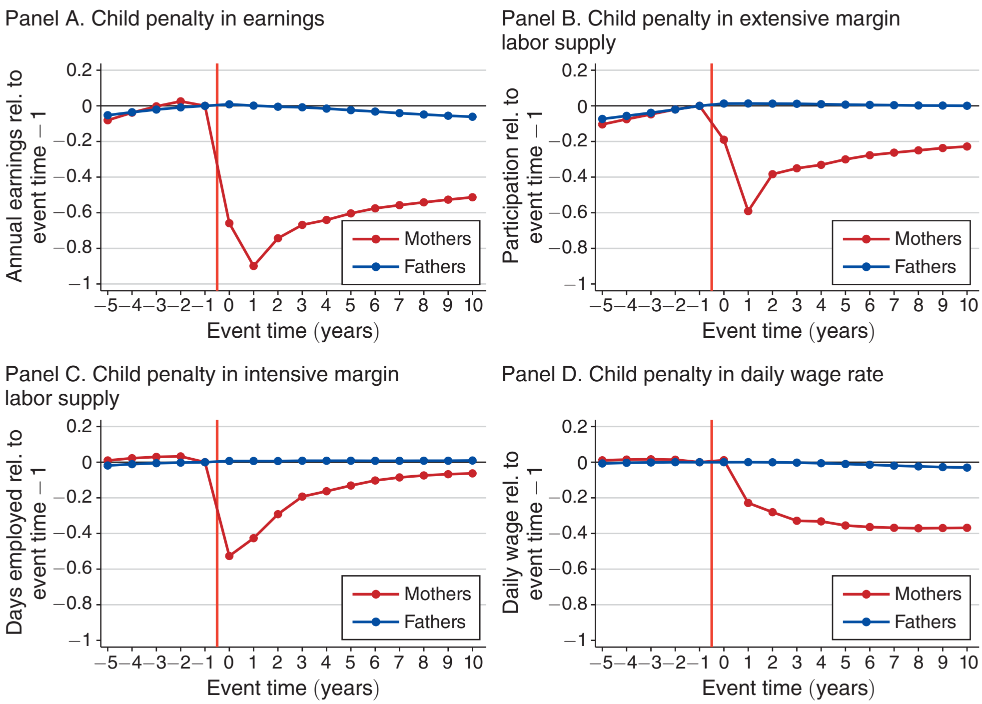

Figure 1. Child Penalties

Notes: The figure shows event time coefficients estimated from equation (2) and normalized as explained in Section IC (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 and 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 euros using the CPI and adjusted for top-coding (see online 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 on having at least one employment spell that year.

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 childbirth correlates strongly with the length of their PL selection. Nine years after childbirth, women in the "30 + 6" scheme exhibit penalties that are about 15 pp 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 PL over the past 60 years.

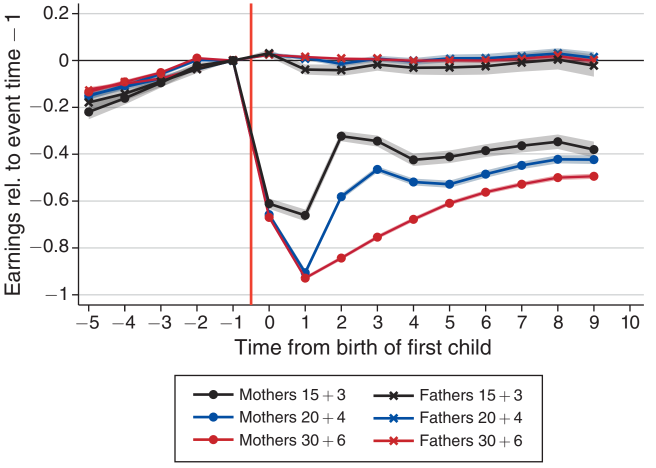

Figure 2. The 2008 Parental Leave Regime: Child Penalty by PL Option

Notes: This figure shows normalized event time coefficients ($P_t^g \equiv \hat{\alpha}_t^g / E[\tilde{Y}_{ist}^g | t]$) 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 PL 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 PL 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 percent confidence intervals on the normalized event time coefficients.

A. *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 PL. 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 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 childbirth.

[17] Before 1990, the PL 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 PL 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 reform increased the maximum duration of PL 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 18 to 30 months.

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 1, 1990, became eligible for 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 1, 1996, and July 1, 2000. This implementation structure lends itself to an RD 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 PL reforms on the full dynamics of female outcomes around the event of the first childbirth.[20] To study this effect, online 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 four years before childbirth 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 childbirth.[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, respectively, 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 PL 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).

[18] Note 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 for 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.

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

[20] We 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.

[21] The 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 PL in that calendar year, their earnings are mechanically lower.

[22] For women who give birth in July 1990, event year 1 corresponds the period going from January to December 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.

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 online 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:

$$(3) \qquad Y_{it}^{w} = \boldsymbol{\alpha}\mathbf{D}_{it}^{Event} + \boldsymbol{\alpha}^{T}\mathbf{D}_{it}^{Event} \cdot Treat_i + \boldsymbol{\alpha}^{B}\mathbf{D}_{it}^{Event} \cdot Birth_i + \boldsymbol{\delta}\mathbf{D}_{it}^{Event} \cdot Treat_i \cdot Birth_i + \boldsymbol{\beta}\mathbf{D}_{it}^{Age} + \nu_{it}.$$

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 1, 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 1 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 $E\left[\tilde{Y}_{it}^{w} \mid t, Birth_i = 1990\right]$ obtained from specification (3) when omitting the contribution of the event dummies. The statistic $\hat{\delta}_t / E\left[\tilde{Y}_{it}^{w} \mid t\right]$ 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 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 percent 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 percent (0.62 percent) 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 online Appendix Figure B.II reveals no sign of a discontinuity at the July 1, 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

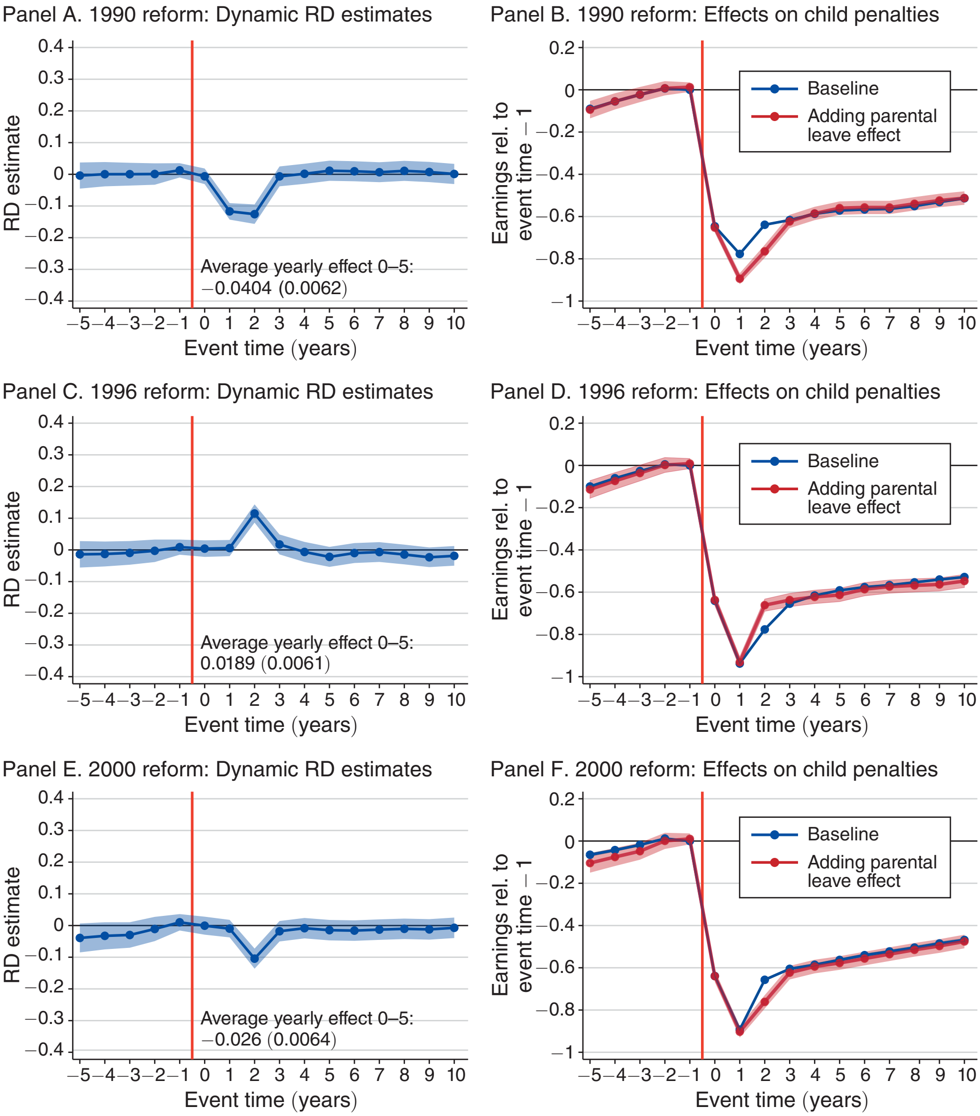

Figure 3. Parental Leave Reforms: Dynamic RD Estimates and Causal Effects on Child Penalties

Notes: The figure reports the estimated effects of the 1990, 1996, and 2000 reforms on child penalties in earnings for women. The left panels plot 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 percent confidence intervals.

event year 2—i.e., in the year where the PL contraction had bite.[23] As before, the reform did not have any long-run effects on the careers of mothers: the increase in

[23] For women who gave birth around the July 1, 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.

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 PL extension from 18 to 30 months. RD plots for each event year are provided in online 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 PL 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 PL by 1 year (but from a lower baseline), we find that the 2000 reform caused a significantly smaller decline of 2.60 percent, as compared to 4.04 percent for the 1990 reform, in the average earnings of mothers over the first 5 years following childbirth.

For completeness, online Appendix Figure B.V considers the effect of PL reform on fathers. Since 1990, fathers have been entitled to six months of PL, 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.[24]

B. *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.[25] 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.[26] 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, online Appendix Figure B.VII shows that take-up of maternity benefits featured a gradual increase starting with children born in January 1960.

[24] Online Appendix Figure B.V further shows for all three reforms that the take-up of paternal leave by fathers does increase. But the magnitudes of both the baseline levels of take-up and absolute increases in take-up are minuscule, explaining why PL reforms did not affect the overall earnings dynamics of fathers.

[25] Maternity benefit payments (Karenzurlaub) technically start at the end of the maternity insurance period (Mutterschutzgesetz), which covers women for eight weeks following delivery.

[26] A notable exception is Schönberg and Ludsteck (2014).

To deal with this issue, instead of the RD approach used above, we analyze the 1961 reform using a 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:

$$(4) \qquad Y_{it}^{w} = \boldsymbol{\alpha} \mathbf{D}_{it}^{Event} + \boldsymbol{\alpha}^{T} \mathbf{D}_{it}^{Event} \cdot T_i + \boldsymbol{\beta} \mathbf{D}_{it}^{Age} + \boldsymbol{\gamma} \mathbf{D}_{it}^{Year} + \nu_{it}.$$

Figure 4 shows the time series of DiD estimates $\hat{\alpha}_t^T / E\left[\tilde{Y}_{it}^{g} | t, T = 1\right]$, i.e., the percentage point impact on child penalties for women who became eligible for 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 PL 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.[27] 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, online Appendix Figure B.VIII 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. In a similar vein, we implemented doughnut RD specifications excluding births in 1960, to flexibly account for trends in outcomes by birth cohorts. We report the results in online Appendix Figure B.IX following a graphical representation similar to Gray et al. (2021) and found results that are extremely close to our original DiD estimates.

C. *Why Does Parental Leave Not Hurt Mothers' Careers?*

Scholars have expressed mixed views on the long-run effects of PL policy on gender inequality. On the one hand, some have voiced concerns that generous PL 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, PL 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 PL fosters female labor market

[27] For 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).

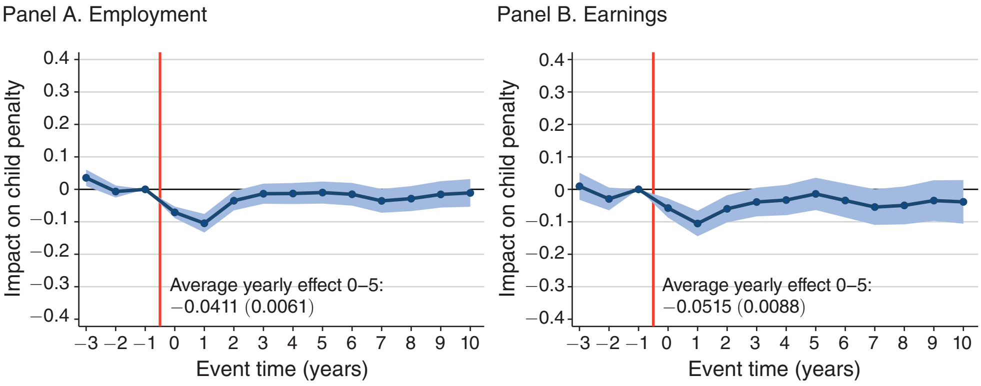

Figure 4. Dynamic Effects of the 1961 Parental Leave Reform

Notes: The figure shows the estimated effects of the introduction of the 1961 reform of PL on the dynamics of female employment (left panel) and female earnings (right panel), based on the REV sample of mothers who give birth to their first child in 1959 or 1961 and are eligible for PL (52 weeks of employment in the two years before birth). In each panel, we report the estimates $\hat{\alpha}_t^T / E\left[\tilde{Y}_{ist}^g \,|\, t, T = 1\right]$ from (4), which correspond to the percentage point change in the child penalty at event time t for women who become eligible for 12 months of maternal leave. Shaded areas correspond to 95 percent confidence intervals.

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 et al. (2016); and Bailey et al. (2019), who find either small or negative effects of PL 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 PL policies over the very long run (about 50 years) and at very different baseline durations (0–3 years). Our analysis show that PL 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.

Nonlinear Effects of Parental Leave.—First, our setting allows us to shed light on nonlinearities in the effects of PL policies, which are usually held responsible for the differences in estimates across contexts and reforms (see Rossin-Slater 2017). One key source of nonlinearity stems from the interaction between the duration of job protection and the duration of cash benefits (Lalive et al. 2014). 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 PL 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

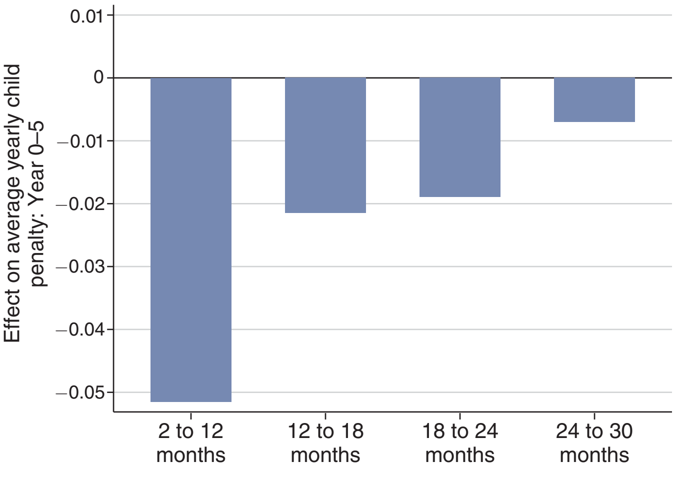

Figure 5. Parental Leave Expansions: Causal Effects by Baseline Duration

Notes: This figure reports the estimated marginal effects of PL extensions at different levels of baseline PL duration on the average yearly child penalty in the first five years following a first childbirth. The three estimates on the right of the graph correspond to the effects of 6 months increments in PL 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 PL, which captures the effect of a 10-month extension (from 2 to 12 months) of paid PL, covered with job protection.

five years following childbirth. The figure shows that the short-run effect of extending PL is highly nonlinear: 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 nonlinearities documented in Figure 5 apply only to the short run, and they are driven entirely by decreasing marginal take-up rates.[28] 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 seems to make any difference to the long-run labor market outcomes of women.

Parental Leave and 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 childcare provision and PL policies drive these results in the Austrian context. But the absence of long-run dynamic effects

[28] Online Appendix Figure B.IV provides direct evidence for this decrease in marginal take-up rates.

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, online Appendix Figure B.X shows that even among mothers in the top quartile of the prebirth 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 PL 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 PL on child penalties and contribute to mask other dynamic effects of the policy.[29] 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 online Appendix Figure B.XI 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.XI 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 2.5 pp on the probability of a second child.[30] However, these fertility responses are unlikely to have much of an impact on child penalties. First, the fertility effects are relatively small—the baseline probability of having a second child is 35 percent within 3 years and 60 percent within 10 years. Second, the relevant fertility outcome is completed fertility, including all parities. In panel C of online Appendix Figure B.XI, 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

[29] Specifically, 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 et al. (2016) find no effect of longer PL on subsequent fertility in the context of Norway. Malkova (2018), on the other hand, finds significant effects on overall fertility in Soviet Russia.

[30] Lalive and Zweimüller (2009) demonstrate that the Austrian PL reforms affected not only 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 PL renewal rule: to qualify for a new PL, mothers giving births to a new child—while still on PL 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 PL period of the first child. In contrast, under the one-year PL duration of the PL regime prior to the 1990 reform, it was difficult to take advantage of the renewal rule for purely biological reasons.

child in an eight-month window around the cutoff date of the 1990 PL reform (panel C). Interestingly, it turns out that any discontinuity at the July 1 cutoff disappears, indicating that eligibility to the PL 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 online Appendix Figure B.XII. We find that accounting for the effects of PL reforms for all subsequent births only slightly increases the impact of PL 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, PL reforms have very small effects on maternal careers.

III. The Impact of Childcare Provision on Child Penalties

Over the last 60 years, as PL generosity was drastically increased in Austria, so was the provision of public childcare. Sixty years ago, nurseries were extremely rare, and the majority of municipalities did not have a kindergarten (preschool). Nowadays, nurseries are common, and kindergartens even more so, with more than 90 percent of all children aged 3–5 enrolled in preschool care. How did this massive childcare expansion affect the labor market outcomes of women relative to men? And did the effects interact with the expansion of PL generosity?

A. *The Expansion of Childcare Provision*

Measuring Childcare Provision.—We start by constructing long time series of childcare supply for each Austrian municipality. The central government collects granular information on all nursery and kindergarten facilities for each municipality and year. The government reports, available since 1988, contain information on the location of each childcare institution, its opening hours, its number of certified teachers and their contracts (part-time versus 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 include 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 childcare coverage for children aged 1–2 (nursery care index) and children aged 3–5 (preschool/kindergarten care index) in each municipality and year:

$$Index\ 1\text{–}2 \;=\; 100 \times \frac{\#FTE\ Childcare\ Spots\ for\ Children\ Age\ 1\text{–}2}{\#Children\ of\ Age\ 1\text{–}2},$$

$$Index\ 3\text{–}5 \;=\; 100 \times \frac{\#FTE\ Childcare\ Spots\ for\ Children\ Age\ 3\text{–}5}{\#Children\ of\ Age\ 3\text{–}5}.$$

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 childcare 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 childcare institution for each municipality and year.[31] We also combine the information from our two indices into an aggregate index of childcare, labeled *Index 1–5*:

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

Online Appendix Figure C.I illustrates the massive expansion of childcare in Austria since the 1980s. It displays the evolution of the nursery care and kindergarten care indices 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 percent. The index increased significantly after that, especially during the 2000s, and is now close to 35 percent. 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 percent in 1988 and then increased sharply through the 1990s and early 2000s to reach virtually full coverage.

B. *The Effect of Local Childcare Expansions on Child Penalties*

To assess the impact of childcare provision on the dynamics of female labor market outcomes, we start by comparing child penalties for women living in municipalities with high versus low childcare coverage. We rank municipalities according to the value of their aggregate childcare 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 PL. To adjust for

[31] Note that we top-code each index at 100, as some communes offer spots in childcare for more than the number of children in their own municipality.

differences in the prebirth characteristics of mothers in different municipalities, we use inverse probability weighting based on prebirth earnings, prebirth employment, and age.[32] 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 5.11 pp lower on average during the first five years following childbirth. This difference is persistent and stable over the longer run. If we scale the difference in earnings penalties by the difference in childcare coverage across the two groups, we obtain a large and significant TOT effect of 0.158 (0.002), suggesting that increasing childcare coverage from 0 to 100 percent would translate into a 16 pp 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 childcare precisely because they want to invest in their careers after becoming mothers. A closer look at the geographical distribution of childcare in 1990 confirms that selection is likely to be an issue. Online Appendix Figure C.II provides a heat map of the childcare index in 1990, showing a stark contrast 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 childcare 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 Childcare Expansions.—While childcare was extremely segmented geographically around 1990, local childcare 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 childcare 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 childcare expansions. Some municipalities expanded childcare in the 1990s, while their neighbors or other geographically similar municipalities expanded childcare later, in the 2000s. Some municipalities saw very steep increases, while others saw smaller and more gradual increases.

The variation in local childcare expansions offers a promising identification strategy. In principle, we are interested in capturing the macro effects of childcare 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 childcare expansion across municipalities

[32] Concretely, we predict, using a probit model, the probability π_i of being in a municipality with above-median childcare 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 childcare 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 $E[1 - \pi_i]/(1 - \hat{\pi}_i)$.

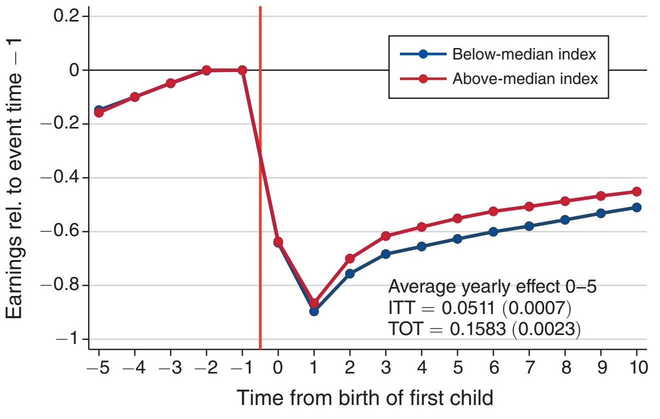

FIGURE 6. CHILD PENALTY BY LEVEL OF CHILDCARE PROVISION INDEX 1–5 IN 1990

Notes: This figure compares child penalties for women who give birth in municipalities with high versus low childcare 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 childcare index (index 1–5) in 1990 and split the sample between municipalities below versus 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 prebirth characteristics of mothers across the two subsamples, we use inverse probability weighting based on prebirth earnings, prebirth employment, and age. The ITT estimate corresponds to the average difference in penalties in the two groups over the first five years after birth. The TOT estimate corresponds to the ITT estimate scaled by the difference in our index 1–5 of childcare coverage across the two groups.

can meaningfully capture macro effects and is therefore appealing from an identification standpoint.[33]

Spatial variation in the expansion of childcare 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 childcare. To deal with such concerns, we isolate episodes of large and sudden increases in childcare provision at the municipal level. Specifically, we focus on events where the nursery care index or the preschool care index of a municipality increases by 20 pp or more in a single year. These events are driven by large supply shocks such as the construction of new facilities or sizable 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 20 pp in their childcare index are assigned to a control group. Our approach is a DiD design in which we compare treated municipalities before and after an episode of childcare expansion to control municipalities before and after a placebo episode.[34] To probe whether treated and control municipalities are balanced on observables, online Appendix Table C.I provides descriptive statistics by treatment status. The

[33] Evidence 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.

[34] To 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.

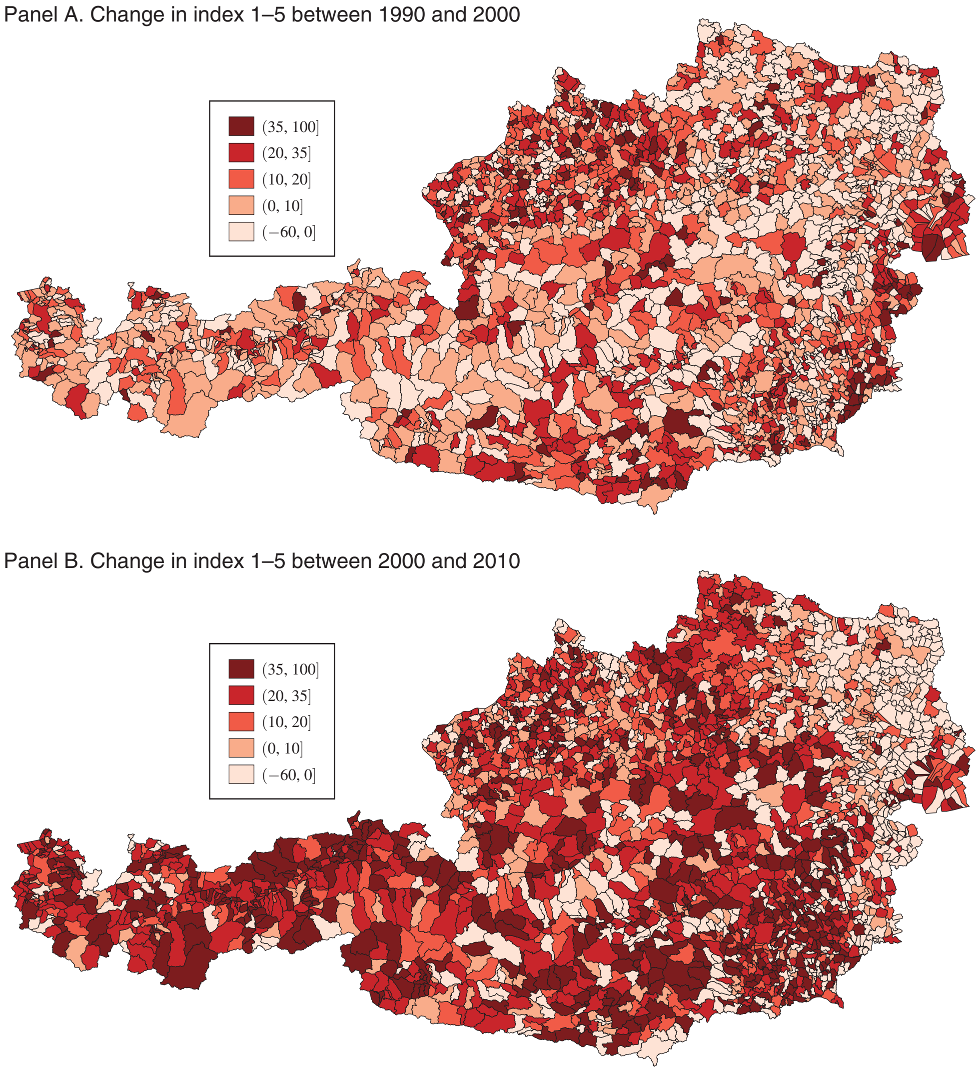

Figure 7. Spatial Variation in Childcare Expansion

Notes: The figure showcases the significant amount of spatial variation in the expansion of local childcare provision in Austria. Each panel corresponds to a heat map that reports for each municipality the change in the value of the childcare 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 childcare provision. See Section IIIA for further details on the construction of the childcare provision index. GIS files of municipality borders provided by Statistik Austria.

table shows that prior to local childcare expansions, the two groups are remarkably similar in terms of population structure and labor market outcomes.

In Figure 8, we document the evolution of childcare provision and childcare 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 40 pp 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 childcare expansions are endogenously driven by the dynamics of local demand. An increase in childcare 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 childcare take-up closely mirrors the dynamics of childcare 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 percent to 28 percent.

Figure 8, panels C and D replicate the analysis for preschool care. These panels reveal similar dynamic patterns as those observed for nursery care. The main difference relates to the baseline level of childcare in treatment versus 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 childcare. In the case of preschool care expansions, treated municipalities are lagging behind and are using the expansions to catch up.[35]

To estimate the effects of childcare expansions, we use the following DiD specification. 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 childcare expansion, indexing the year of the expansion event to $s = 0$. Because we are interested in mapping out the effects of childcare on the full dynamics of labor market outcomes around childbirth, 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:

$$(5) \qquad Y_{ist} = \boldsymbol{\alpha}_t \mathbf{D}_{ist}^{Exp} + \boldsymbol{\alpha}_t^T \mathbf{D}_{ist}^{Exp} \cdot T_i + \boldsymbol{\beta} \mathbf{D}_{ist}^{Age} + \boldsymbol{\gamma} \mathbf{D}_{ist}^{Year} + \nu_{ist},$$

[35] We also investigated the presence of local treatment spillovers that may arise if mothers in neighboring control municipalities take their children to treated municipalities for care so that they could increase their attachment to work. To this effect, we constructed a distance matrix of driving times between municipalities. Online Appendix Figure C.III shows the fraction of children age 1–2 in nursery care (panel A) and the fraction of children age 3 to 5 in preschool care (panel B) for treatment municipalities and for control municipalities within ten minutes of driving time of the treated municipalities and, finally, for control municipalities more than ten minutes away from treated municipalities. Interestingly, the fraction of children from neighboring municipalities in institutional childcare displays (i) a small increase on impact (i.e., in the year of childcare expansion in treated municipalities) and (ii) a slightly stronger trend upward after the expansion for nursery care take-up. While this is evidence of the presence of some local spillovers, their magnitude is very modest, and our results are virtually unchanged when excluding these neighboring municipalities from our control group.

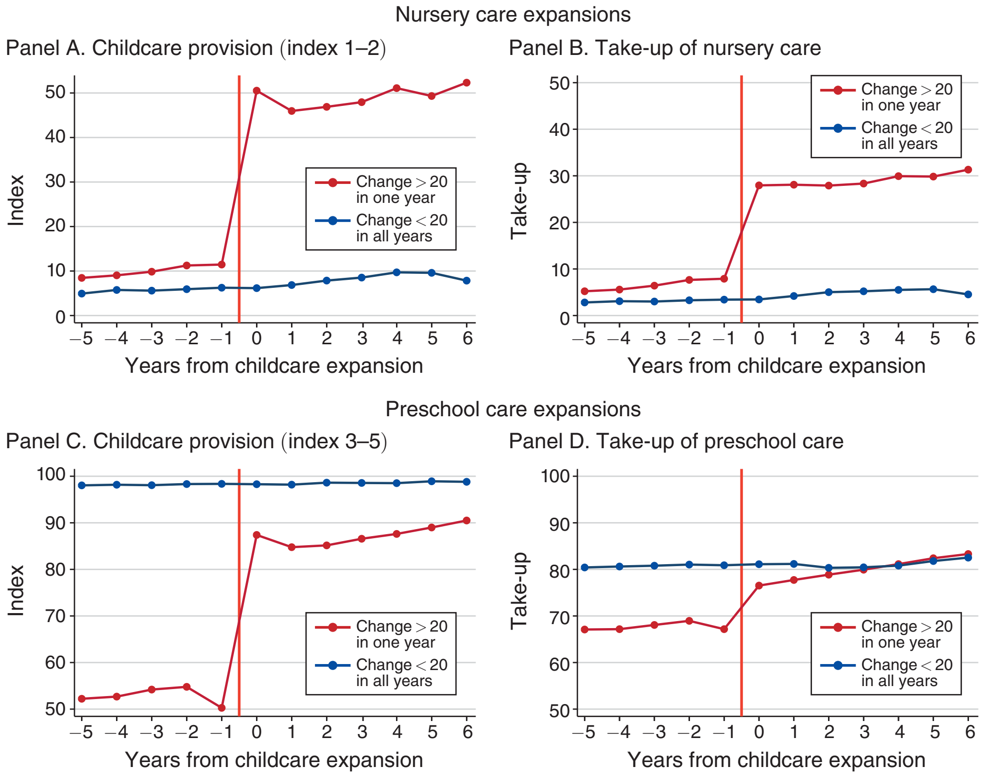

FIGURE 8. LOCAL CHILDCARE SHOCKS: EVOLUTION OF CHILDCARE PROVISION AND TAKE-UP FOR TREATED AND CONTROL MUNICIPALITIES

Notes: The figure shows the evolution of childcare 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 childcare provision. A nursery care (resp. preschool care) expansion event is defined as an increase of more than 20 points in the municipal nursery care (resp. preschool 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.

where $\mathbf{D}^{Exp}$ is a vector of event time dummies with respect to childcare expansion. As before, we control for calendar year fixed effects and age fixed effects. Identification relies on the usual parallel trend assumption that absent childcare 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 childbirth is stacked into a single specification. That is, we estimate

$$Y_{ist} = \sum_j \left(\alpha_j \mathbf{D}_{isj}^{Exp}\right) \cdot \mathbf{1}\{j = t\} + \sum_j \left(\alpha_j^T \mathbf{D}_{isj}^{Exp} \cdot T_i\right) \cdot \mathbf{1}\{j = t\} + \beta \mathbf{D}_{ist}^{Age} + \gamma \mathbf{D}_{ist}^{Year} + \nu_{ist}. \tag{6}$$

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

$E\left[\tilde{Y}_{ist}\,|\,t\right]$. To get a visual representation of our DiD estimates, we plot, for each t, the standardized coefficients $\hat{\alpha}_{st}/E\left[\tilde{Y}_{ist}\,|\,t\right]$ and $\left(\hat{\alpha}_{st} + \hat{\alpha}_{st}^{T}\right)/E\left[\tilde{Y}_{ist}\,|\,t\right]$ 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 childcare 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 childcare expansion. To account for spatial correlation in unobserved labor market shocks, we cluster standard errors at the municipal level.

The Limited Effects of Childcare Expansions.—In Figure 9, we present our results for nursery care expansions. Figure 9, panel A considers women one year before childbirth ($t = -1$) 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 childcare did not lead to anticipatory effects among women about to have children.

Figure 9, panel B considers women one to two years after childbirth, corresponding to the time where children are enrolled in nursery care. As in Figure 9, 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 1.1 pp (standard error 1.4 pp). If we scale this estimate by the change in nursery care index, we find that increasing municipal nursery care coverage from 0 to 100 percent would reduce the child penalty by only 3.1 pp (standard error 4.0 pp).

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 three years before birth to five years after birth. The results, reported in online Appendix Figure C.IV, 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 five years following the birth of a child. These estimates imply that an increase in coverage from 0 to 100 percent would reduce the child penalty by only 3.0 pp (standard error 4.3 pp) in the first 5 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

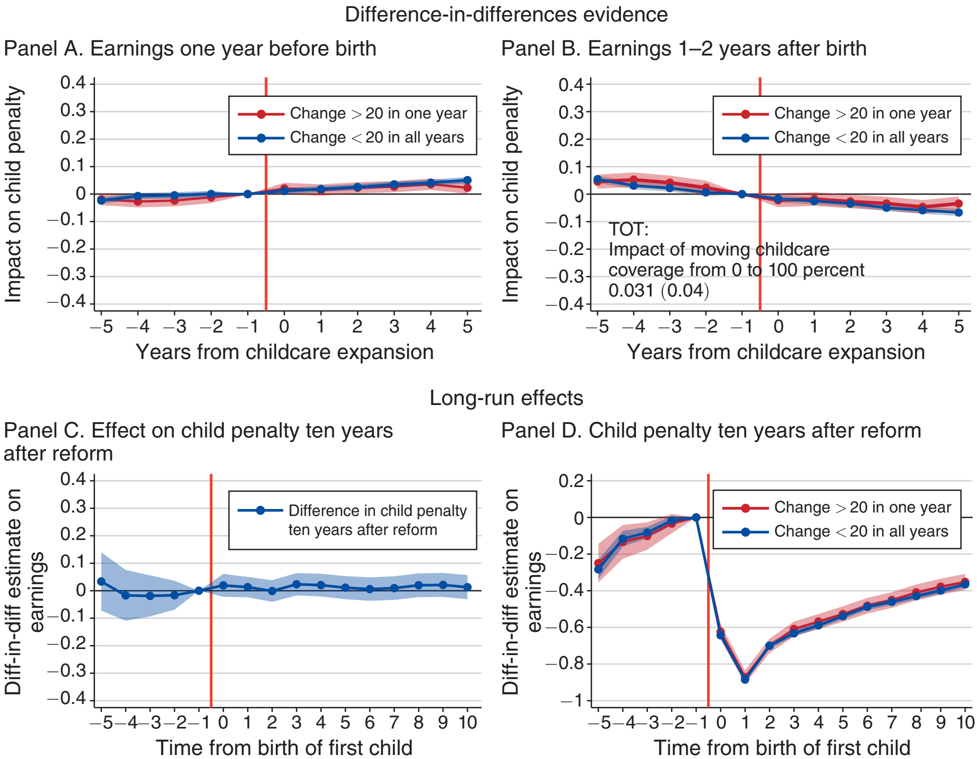

Figure 9. Estimated Effects of Nursery Care Expansions

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 versus 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 $\hat{\alpha}_{st}^{T}$ 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 $\hat{\alpha}_{st}^{T}$ for earnings at all event times t with respect to birth of the first child, evaluated at $s = 10$, that is, ten 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 PL 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 percent confidence intervals.

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 PL 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 childcare effects is the availability of long maternal leaves during our period of analysis. Mothers can enroll 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 preschool years—i.e., when children are between three and five years old. Yet, Figure 10 shows that episodes of large preschool expansions had no significant impact on child penalties either. Panels A and B show that earnings, either just before birth or three to five 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 preschool care.[36]

C. *Interaction between Childcare and Parental Leave Policies*

Given that both childcare and PL 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, online Appendix Figure C.VIII shows the effects of the 1990 PL extension, comparing women living in municipalities with high and low levels of childcare coverage. We first split municipalities by the value of their aggregate index of nursery and preschool 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 childcare provision. Panel B focuses specifically on the supply of nursery care, while panel C focuses on the supply of preschool care: in both cases, we do not detect any significant differences in the impact of the PL reform between municipalities with high and low provision of childcare.

D. *Why Does Childcare Not Improve Mothers' Careers?*

The existing literature on the impact of childcare policies on mothers' labor market outcomes has produced mixed results. Some studies estimate positive and sometimes large effects of childcare subsidies on female labor supply (e.g., Baker, Gruber, and Milligan 2008; Berlinski and Galiani 2007; Cascio 2009; 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 childcare provision to improve female labor supply. Studies from the United States (Fitzpatrick 2010, 2012; Barua 2014), Norway (Havnes and Mogstad 2011a), France (Goux and Maurin 2010), and Spain (Nollenberger and Rodríguez-Planas 2015) find that childcare 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 childcare in Austria shows a precisely estimated zero effect on mothers' careers, in the short run as well as in the long run.

[36] For completeness, online Appendix Figure C.VI shows the effects of childcare expansions on female employment, following specification (6). Results confirm that neither nursery care expansions nor preschool care expansions had any significant effects on employment. We also consider, in online Appendix Figure C.VII, the effects of childcare expansions on the earnings of fathers, and do not detect any significant effect.

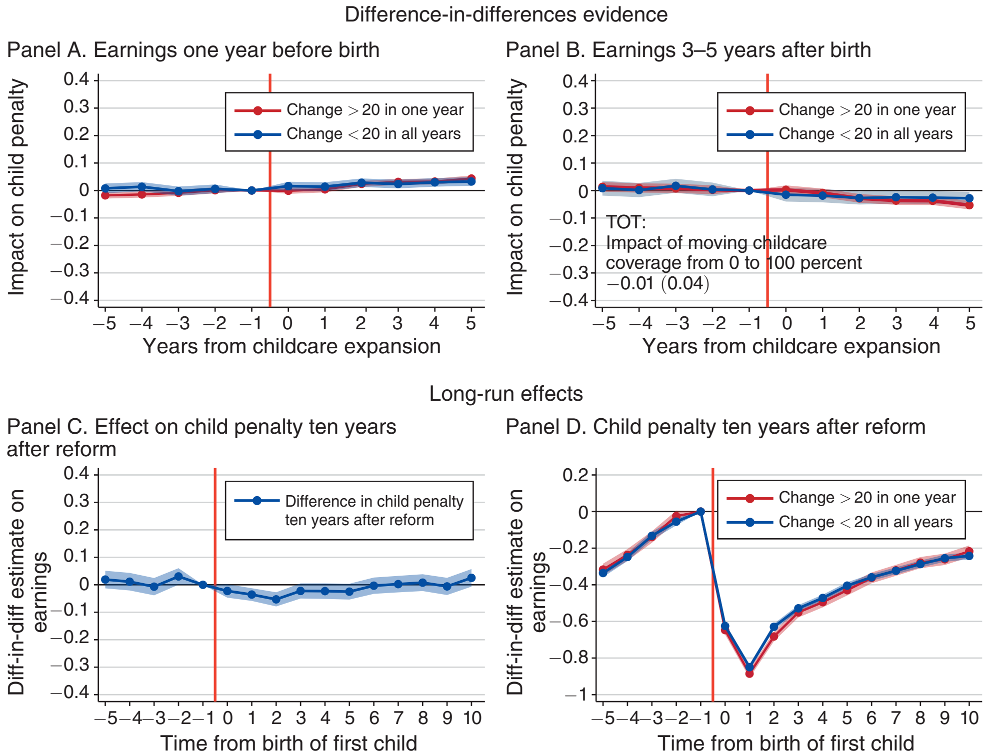

FIGURE 10. ESTIMATED EFFECTS OF PRESCHOOL CARE EXPANSIONS

Notes: The figure replicates the analysis of Figure 9 for preschool care expansions events. It shows the DiD estimates of the effects of preschool 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 versus control municipalities, around the event of a preschool 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 preschool care index. In panel C, we plot the DiD estimate $\hat{\alpha}_{st}^{T}$ for earnings at all event times t with respect to birth of the first child, evaluated at $s = 10$, that is, ten years after the local preschool care expansion. Panel C is therefore the mirror image, for preschool care reforms, of panels A, C, and E of Figure 3 in the case of PL reforms. Based on the estimates of panel C, we draw in panel D counterfactual child penalties under the pre- and post-preschool care expansion regimes. Standard errors are clustered at the municipal level, and shaded areas correspond to 95 percent confidence intervals.

But why are large increases in the take-up of childcare 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 childcare may simply crowd out informal childcare arrangements such as care by relatives. Second, mothers may face large resource and utility costs of increasing labor supply even after childcare 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 by the Austrian microcensus in 1995 and 2002. These surveys contain household-level information on time use, preferences, and childcare arrangements. To investigate the correlation between these outcomes and childcare provision, we match the survey information to our measures of institutional childcare provision at the political district level.[37]

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 childcare index (index 1–5).

The results are presented in Figure 11. We first consider the relationship between the childcare index and the survey measures of childcare take-up (panel A) and maternal employment (panel B). For these two outcomes, the correlation with the childcare index can be directly compared to the causal estimates from our local expansion design. The results in panels A and B show that the correlations closely match our quasi-experimental estimates. For take-up, we find that a 1 percentage point increase in the childcare index translates into a 0.36 (0.09) 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 preschool expansions together, we find a take-up rate of 0.35. Similarly, both the cross-sectional and local expansion designs are congruent in finding no significant correlation between childcare coverage and mothers' probability of employment.

Having established the internal validity of the cross-sectional design, we turn to the correlation between institutional childcare provision and households' childcare arrangements. Panel C considers weekly hours spent on maternal childcare. The panel shows that the overall amount of maternal care is very high, with mothers spending about 40 hours per week on it 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 childcare in districts where the childcare index is below 30 percent and in districts where the index is above 70 percent. While higher childcare 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 0.35 (0.11) percentage points when the childcare index increases by 1 percentage point. This suggests that working mothers use relatives as childcare providers prior to the availability of publicly provided childcare and then switch to nurseries and preschools when they become available.[38]

[37] The 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.

[38] In a recent study for Austria, Frimmel et al. (2022) 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 that is particularly strong when grandmothers live close to the grandchild.

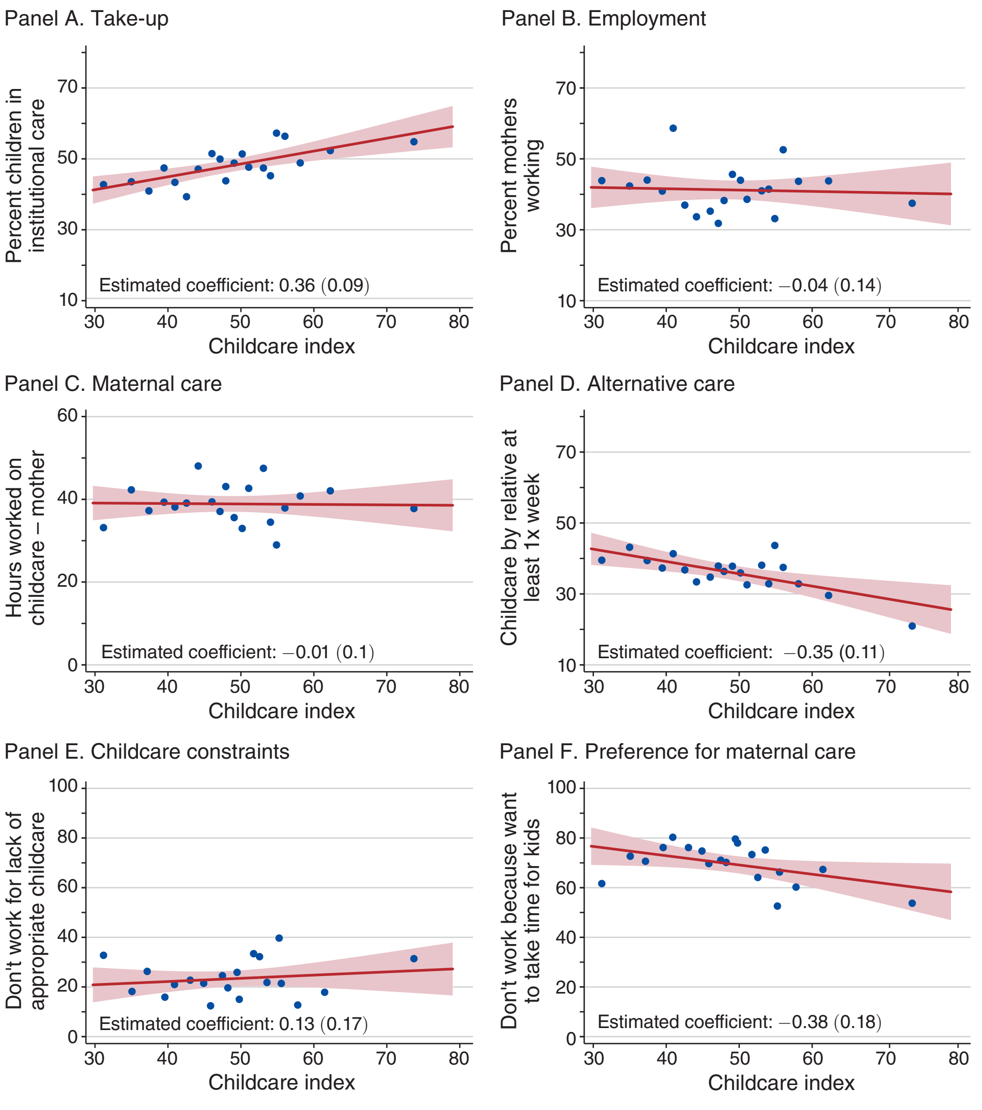

Figure 11. Reported Childcare Arrangements and Preferences: Residual Correlations with Childcare Provision

Notes: The figure reports correlations between measures of institutional childcare provision at the political district level and household-level information on time use, preferences, and childcare 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 childcare 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 childcare. 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 childcare, 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 childcare, 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.

But what about women who do not work: are they facing constraints that prevent them from increasing labor supply? In panel E, we correlate childcare provision with the fraction of mothers reporting that their career is constrained by a lack of appropriate childcare. 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 childcare provision than it is in districts with high levels.

If childcare 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 percent) 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 childcare provision. This indicates that expanding access to institutional childcare 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 childcare provision in Austria has had no effect on female labor market outcomes.

Fertility Effects.—To fully understand the impact of childcare expansions on female careers, we finally need to account for the possibility that the expansions affected fertility. If the increased availability of childcare induced women to have more children, this may have mechanically increased child penalties, all else being equal. That is, the null effect of childcare 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 childcare in Austria over the last 60 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 childcare provision on fertility.

The existing literature offers relatively little guidance on this question. Few studies provide clear evidence on the fertility effects of childcare provision.[39] Fortunately, our setting allows for estimating the effect of municipal childcare 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. Online Appendix Figure C.IX plots the evolution of fertility rates in treated and control municipalities around the time of large shocks in childcare supply. Looking at the long-run effects, ten years after the expansion, we find small and statistically insignificant effects of childcare expansions on fertility. This is true for both nursery and preschool care. These findings suggest that our null effect of childcare expansions are not the result of offsetting fertility effects.

[39] An exception is Bauernschuster, Hener, and Rainer (2016), who find significant fertility effects of local childcare expansions in Germany.

V. 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 60 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.

A. *Child-Related Gender Inequality over the Last 60 Years*

The consistency and quality of the Austrian administrative data allow 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).[40] This allows us to measure, for all years s between 1953 and 2017, the total gender gap in raw earnings, $\Delta_s \equiv 1 - E\left[Y_{ist}^w \mid s\right]/E\left[Y_{ist}^m \mid s\right]$.

For each year, we 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 Søgaard (2019). This strategy relies on using child penalty estimates to get a measure of the counterfactual gender gap, $\tilde{\Delta}_s \equiv 1 - E\left[\tilde{Y}_{ist}^w \mid s\right]/E\left[Y_{ist}^m \mid s\right]$. 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:

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

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 $E\left[\tilde{Y}_{ist}^w \mid s, t\right]$.

[40] See Christl and Wagner (1981). We provide all the details of the data sources and construction of our gender gap measures in online Appendix D.1. 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. Second, we account for the extensive margin of labor supply by including individuals with zero earnings. In that sense, our measure of the gender gap encompasses the broadest possible set of behaviors underlying labor market outcomes.

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 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:

$$(8) \qquad Y_{ist}^{w} = \sum_{y} \left(\boldsymbol{\alpha}_{y}^{w} \mathbf{D}_{ist}^{Event} \right) \cdot \mathbf{1}\{y = s\} + \boldsymbol{\beta}^{w} \mathbf{D}_{ist}^{Age} + \boldsymbol{\gamma}^{w} \mathbf{D}_{ist}^{Year} + \nu_{ist}^{w}.$$

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 $\hat{\alpha}_{st}^{w}$.[41] Overall, results reported in online 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.

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 30 pp over the last 60 years, the gender gap in earnings remains very large in Austria: the average woman earns 40 percent 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 percent less than men *because of children*. This reflects the large child penalties experienced by Austrian mothers. Interestingly, child-related gender inequality has been quite stable over the past six decades.[42] 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 all 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 Søgaard (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.

[41] Note 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 postevent 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 online Appendix D.

[42] This is despite long-run penalties, from 10 to 30 years after birth, having declined significantly, as we showed above. The reason for this apparently counterintuitive result is that penalties are expressed in percentage of counterfactual earnings of mothers. And the lower long-run penalties experienced by Austrian mothers from more recent cohorts apply to a much larger base, as the earnings of mothers and fathers prior to birth have converged significantly over the same period.

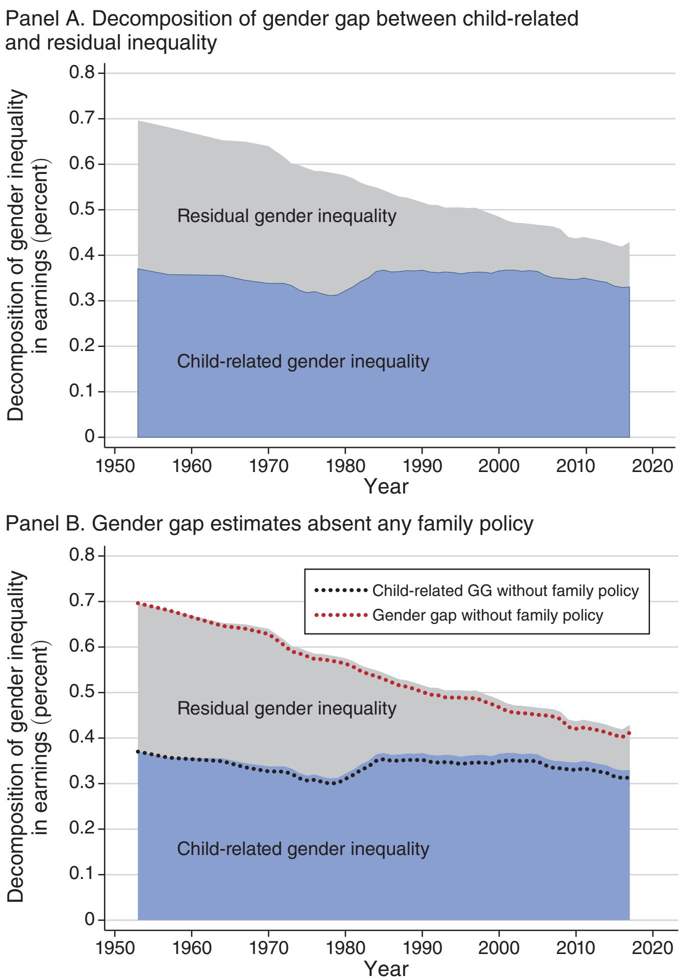

Figure 12. Gender Gap Decomposition and Counterfactual Inequality absent Family Policies

Notes: This figure decomposes the gender gap (GG) 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 gray shaded areas) into child-related (blue) and residual factors (gray), following the methodology of Kleven, Landais, and Søgaard (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 PL duration and our childcare 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 II and III. See text for details.

B. *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 preschool 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, $d\Delta_s / d\tau_k$. We can express this as a function of the causal effect on the dynamics of child penalties, $dP_t/d\tau_k$ at each t. That is, we have

$$\text{(9)} \qquad \frac{d\Delta_s}{d\tau_k} = -\frac{dE\left[Y_{ist}^w \mid s\right]/d\tau_k}{E\left[Y_{ist}^m \mid s\right]}$$

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

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 $d\psi_{st}/d\tau_k = 0$.[43] 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 $\left(\tau_{PL}, \tau_N, \tau_{PS}\right)$. However, we can leverage our empirical finding of no interaction effects between PL and childcare 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., $\left(dP_t/d\tau_k\right)\left(\tau_k\right)$. This, in turn, allows us to easily explore counterfactual scenarios where we change all three policies together.

The scenarios we study involve large, nonmarginal 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., $\left(dP_t/d\tau_k\right)\left(\tau_k\right)$ for all event times t. The second is that we can account for nonlinearity in the effects of these policies. For PL policies, we have identified the effect of changing the duration of PL benefits at different baseline durations, starting from no PL at all. For childcare provision, our estimates show that even massive reforms—increases in the level of coverage of more than 40 pp—do not have any significant effects on child penalties in both the short and long runs. This suggests that nonlinearities are not important for the provision of childcare.

Gender Inequality without Parental Leave and Institutional Childcare.—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 PL reforms and setting the level of nursery and preschool 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 policies over the last 60 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 percent absent family policies, compared to the observed gap of 43.7 percent. The gender gap caused by children would have been 30.1 percent instead of the observed gap of 34.7 percent.

[43] We explore in online Appendix D.3 scenarios in which we relax this assumption and allow family policies to affect fertility as well.

To provide additional insight, online 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 PL expansions, assuming that childcare provision follows the same path as actually observed. The graph highlights that PL expansions on their own have increased gender inequality, although the effect is small. In the absence of PL, we find that gender inequality in earnings would have been about 3 pp smaller today. This result stems from the negative impact of PL policies on earnings in the short run. The absence of long-run effects on the careers on women renders the total impact of PL policies on gender inequality modest. While PL policies do have some effect on the evolution of the gender gap, panels B and C show that childcare policies have had absolutely no effect.

In online Appendix D.3, we provide further details regarding the robustness of our decomposition and counterfactual policy analysis. We first show that our null results are precise: we compute conservative upper and lower bounds on counterfactual gender gaps and show that we can rule out large positive effects of family policies on gender inequality. Second, we also provide evidence of the robustness of our analysis to equilibrium effects, to anticipatory responses prior to childbirth, and to fertility responses.

V. 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 60 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 effects of all family policy variation, our approach is robust to policy interactions and other nonlinearities that normally provides threats to out-of-sample extrapolations of marginal treatment effects.[44]

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 US has not gone as far as Europe, proposals to expand family policies are the subject of ongoing debate. We find that PL 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 childcare has had no effect on the secular decline in gender inequality. This stands in contrast to the strong cross-country correlation between childcare subsidization and gender gaps, while within-country quasi-experimental studies of specific reforms have led to mixed results.

[44] 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.

Table 1—Gender Gap Decomposition under Observed and Counterfactual Family Policy Regimes

Year	1953	1964	1970	1980	1990	2000	2010
Panel 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
Decomposition between child-related and residual inequality							
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
Panel B. Counterfactual family policy scenario							
No PL and no childcare (baseline)							
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
No PL and no childcare (optimistic)							
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
No PL and no childcare (pessimistic)							
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 euros 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" excercise uses, instead of our point estimates, the upper bound of the 95 percent confidence interval on $\left(dP_t/d\tau_k\right)\left(\tau_k\right)$ for each policy change. The second, or "pessimistic," scenario uses the lower bound of the 95 percent confidence interval. See text for details.

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 et al. 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

Andresen, Martin Eckhoff, and Tarjei Havnes. 2018. "Child Care, Parental Labor Supply and Tax Revenue." IZA Discussion Paper 11576.

Baierl, Andreas, and Markus Kaindl. 2011. "Kinderbetreuung in Österreich: Rechtliche Bestimmungen und die Reale Betreuungssituation." Austrian Institute for Family Studies Working Paper 77.

Bailey, Martha J., Tanya S. Byker, Elena Patel, and Shanthi Ramnath. 2019. "The Long-Term Effects of California's 2004 Paid Family Leave Act on Women's Careers: Evidence from US Tax Data." NBER Working Paper 26416.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4): 709–45.

Barua, Rashmi. 2014. "Intertemporal Substitution in Maternal Labor Supply: Evidence Using State School Entrance Age Laws." *Labour Economics* 31: 129–40.

Bauernschuster, Stefan, Timo Hener, and Helmut 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, Charles L., II., and Christopher J. Ruhm. 2016. "The Effects of Paid Family Leave in California on Labor Market Outcomes." *Journal of Policy Analysis and Management* 35 (2): 333–56.

Berlinski, Samuel, and Sebastian Galiani. 2007. "The Effect of a Large Expansion of Pre-Primary School Facilities on Preschool Attendance and Maternal Employment." *Labour Economics* 14 (3): 665–80.

Bertrand, Marianne. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1543–590. Amsterdam, Elsevier.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3): 228–55.

Blau, Francine D., and Lawrence M. Kahn. 2013. "Female Labor Supply: Why Is the United States Falling Behind?" *American Economic Review: Papers and Proceedings* 103 (3): 251–56.

Boelmann, Barbara, Anna Raute, and Uta Schönberg. 2020. "Wind of Change? Cultural Determinants of Maternal Labor Supply." Queen Mary University of London School of Economics and Finance Working Paper 914.

Cascio, Elizabeth U. 2009. "Maternal Labour Supply and the Introduction of Kindergartens into American Public Schools." *Journal of Human Resources* 44 (1): 140–70.

Christl, Josef, and Michael Wagner. 1981. "Geschlechtsspezifische Einkommensunterschiede: Österreich 1953–1979; i.a. des BM fuer Sozial Verwaltung." Vienna: Institut für Höhere Studien.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy* 126 (6): 2356–409.

Cortés, Patricia, and Jessica Pan. 2020. "Children and the Remaining Gender Gaps in the Labor Market." NBER Working Paper 27980.

Dahl, Gordon B., Katrine V. Løken, Magne Mogstad, and Kari Vea Salvanes. 2016. "What Is the Case for Paid Maternity Leave?" *Review of Economics and Statistics* 98 (4): 655–70.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91 (4): 795–813.

Ejrnæs, Mette, and Astrid Kunze. 2013. "Work and Wage Dynamics around Childbirth." *Scandinavian Journal of Economics* 115 (3): 856–77.

Fitzpatrick, Maria Donovan. 2010. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten." *Journal of Labor Economics* 28 (1): 51–85.

Fitzpatrick, Maria Donovan. 2012. "Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool." *Journal of Human Resources* 47 (3): 583–612.

Frimmel, Wolfgang, Martin Halla, Bernhard Schmidpeter, and Rudolf Winter-Ebmer. 2022. "Grandmothers' Labor Supply." *Journal of Human Resources* 57 (5): 1645–89.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–119.

Goldin, Claudia, and Lawrence F. Katz. 2016. "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation." *Journal of Labor Economics* 34 (3): 705–46.

Goux, Dominique, and Eric Maurin. 2010. "Public School Availability for Two-Year Olds and Mothers' Labour Supply." *Labour Economics* 17 (6): 951–62.

Gray, Colin, Adam Leive, Elena Prager, Kelsey B. Pukelis, and Mary Zaki. 2021. "Employed in a SNAP? The Impact of Work Requirements on Program Participation and Labor Supply." NBER Working Paper 28877.

Havnes, Tarjei, and Magne Mogstad. 2011. "Money for Nothing? Universal Child Care and Maternal Employment." *Journal of Public Economics* 95 (11-12): 1455–65.

Kleven, Henrik Jacobsen. 2014. "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives* 28 (4): 77–98.

Kleven, Henrik, and Camille Landais. 2017. "Gender Inequality and Economic Development: Fertility, Education, and Norms." *Economica* 84 (334): 180–209.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2021. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." American Economic Review: Insights 3 (2): 183–98.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109: 122–26.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2024. "Replication Data for: Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E186541V1.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2018. "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" NBER Working Paper 24740.

Lalive, Rafael, and Josef Zweimüller. 2009. "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *Quarterly Journal of Economics* 124 (3): 1363–402.

Lalive, Rafael, Camille Landais, and Josef Zweimüller. 2015. "Market Externalities of Large Unemployment Insurance Extension Programs." *American Economic Review* 105 (12): 3564–96.

Lalive, Rafael, Analía Schlosser, Andreas Steinhauer, and Josef Zweimüller. 2014. "Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits." *Review of Economic Studies* 81 (1): 219–65.

Malkova, Olga. 2018. "Can Maternity Benefits Have Long-Term Effects on Childbearing? Evidence from Soviet Russia." *Review of Economics and Statistics* 100 (4): 691–703.

Nollenberger, Natalia, and Núria Rodríguez-Planas. 2015. "Full-Time Universal Childcare in a Context of Low Maternal Employment: Quasi-Experimental Evidence from Spain." *Labour Economics* 36: 124–136.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics* 8: 405–34.

Olivetti, Claudia, and Barbara Petrongolo. 2017. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives* 31 (1): 205–30.

Rossin-Slater, Maya. 2017. "Maternity and Family Leave Policy." NBER Working Paper 23069.

Rossin-Slater, Maya, Christopher J. Ruhm, and Jane 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): 224–45.

Ruhm, Christopher J. 1998. "The Economic Consequences of Parental Leave Mandates: Lessons from Europe." *Quarterly Journal of Economics* 113 (1): 285–317.

Schönberg, Uta, and Johannes Ludsteck. 2014. "Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth." *Journal of Labor Economics* 32 (3): 469–505.

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi. 2009. "Austrian Social Security Database." Institute for Empirical Research in Economics Working Paper 410.