DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

The Evolution of Child-Related Gender Inequality in Germany and The Role of Family Policies, 1960-2018

by

Ulrich GLOGOWSKY
Emanuel HANSEN
Dominik SACHS
Holger LÜTHEN

Working Paper No. 2408
September 2024

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

ulrich.glogowsky@jku.at

The Evolution of Child-Related Gender Inequality in Germany and The Role of Family Policies, 1960-2018

Ulrich Glogowsky, Emanuel Hansen, Dominik Sachs, Holger Lüthen*

September 17, 2024

Abstract

Using German administrative data from the 1960s onward, this paper (i) examines the long-term evolution of child-related gender inequality in earnings and (ii) assesses the impact of family policies on this inequality. We present three sets of findings. First, child penalties (i.e., the percentage of potential earnings lost due to children) have strongly increased over the last decades. Mothers who had their first child in the 1960s faced much smaller penalties than those who gave birth in the 2000s. Second, we decompose overall gender inequality into child-related and child-unrelated components. Over our sample period, the fraction of overall inequality attributed to children rose from 14% to 64%. This trend not only resulted from the growing child penalties but also from rising potential earnings of mothers. Intuitively, in later decades, mothers had more income to lose from child-related career breaks. Third, we investigate the role of policy decisions in this rise in child penalties. Parental leave expansions between 1979 and 1992 amplified child penalties and contributed nearly one-third to the increase in child-related gender inequality. Instead, a parental benefit reform in 2007 mitigated further increases. While the third set of results highlights the role of family policies, the first two imply that sidelining mothers becomes increasingly costly over time.

*Glogowsky: Johannes Kepler University Linz and CESifo (email: ulrich.glogowsky@jku.at). Hansen: Ludwig-Maximilians-University of Munich and CESifo (email: emanuel.hansen@econ.lmu.de). Sachs: University of St. Gallen and CESifo (email: dominik.sachs@unisg.ch). Lüthen: German Federal Ministry for Economic Affairs and Climate Action (email: holger.luethen@bmwk.bund.de). Special thanks go to Timm Bönke for sharing resources and exchanging ideas. For helpful discussions and comments, we also thank Barbara Boelmann, Pierre Boyer, Thiess Bütter, Lisa Dettling, Christina Felfe, Martin Halla, Andreas Haufler, Paul Hufe, Han Ye, Lavinia Kinne, Tommy Krieger, Wojciech Kopczuk, Andreas Peichl, Nicole Schneeweis, Sebastian Siegloch, Jakob Søgaard, Anna Raute, Johannes Rincke, Stefanie Stantcheva, Holger Stichnoth, Michèle Tertilt, Arne Uhlendorf, and Luisa Wallossek. Marvin Immesberger, Rafael Pfeiffer, and Helene Wagner provided valuable research assistance. Finally, we are also thankful for the comments and suggestions received from conference and seminar participants on many occasions, including those at the workshop "Capital vs. Labor" in St. Gallen. Financial support from Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged. The authors have no conflicts of interest.

1 Introduction

A growing body of literature documents substantial child penalties in earnings for mothers [Angelov *et al.*, 2016; Kleven *et al.*, 2019a,b]: Due to parenthood, they experience substantial earnings losses. By contrast, parenthood does not affect fathers' long-term earnings. These unequal impacts of parenthood on fathers' and mothers' careers translate into substantial child-related gender inequality in earnings. Researchers have documented these patterns across numerous countries [Kleven *et al.*, 2019a; Cortés and Pan, 2023; Kleven *et al.*, 2024a]. However, most available papers focus on recent decades.[1] Furthermore, the debate on the forces driving child-related gender inequality remains inconclusive despite recent progress [Andresen and Nix, 2023; Gruber *et al.*, 2023; Huber and Rolvering, 2023; Kleven, 2023; Kleven *et al.*, 2024a,b; Krapf *et al.*, 2020].[2]

This paper analyzes the evolution and determinants of child-related earnings inequality in Germany over more than fifty years. Our study contributes to the literature by presenting three sets of results. First, we show that child penalties — the percentage of potential earnings mothers lose due to having children — have grown substantially over the past decades. Mothers giving birth in the 1960s faced much smaller penalties than those in the 2000s. Second, to explore whether and how this increase in child penalties led to a rise in gender inequality, we decompose overall gender inequality into child-related and child-unrelated components. Over our sample period, the fraction of inequality attributed to children grew from 14% to 64%. We clarify that two factors explain this trend: Mothers in more recent decades not only faced higher child penalties but also had more potential earnings at stake (i.e., they would have earned more without children). Third, we show that six parental leave reforms contributed to this surge in child-related inequality. A parental benefit reform, instead, lowered inequality.

Data and empirical approaches. Our setting offers rich administrative data and clean identifying variation that jointly enable us to establish these results. The basis

[1] Two points are of note. First, the extent of child penalties and their impact on gender inequality vary by context [Angelov *et al.*, 2016; Kleven *et al.*, 2019a; Andresen and Nix, 2022; Cortés and Pan, 2023]. Second, there are a few papers that also study earlier decades [see, e.g., Kleven *et al.*, 2024a,b].

[2] A growing number of studies indicate that norms play a critical role in shaping child penalties [Kleven *et al.*, 2019a,b; Kleven, 2023]. The amount of inequality linked to child penalties also appears to vary systematically with economic development and measures for structural change [Kleven *et al.*, 2024a]. There is no consensus on the role of family policy, by contrast. Some studies find limited effects of various family policies [Andresen and Nix, 2023; Kleven *et al.*, 2024b], while others come to different conclusions [Gruber *et al.*, 2023; Huber and Rolvering, 2023; Krapf *et al.*, 2020].

of our analyses is monthly pension register data (1960 − 2018). These data contain information on the complete earnings trajectories of West and East German mothers born after 1934. The combination of a long observation period and the granularity of monthly data allows us to precisely track how the careers of many cohorts of mothers evolved around childbirth. We apply two empirical approaches to these data to derive our results. First, we estimate child penalties following the birth of the first child using the event study approach proposed by Kleven *et al.* [2019a,b]. The key benefit of this method is that it allows us to estimate child penalties for the population of all mothers.[3] Second, we exploit dynamic regression discontinuity designs to estimate the effects of six parental leave reforms and one parental benefit reform on child penalties and gender inequality. A common feature of all these reforms was that the new rules only applied to parents who delivered after specific birthdate cutoffs. Thus, we can compare the earnings paths of mothers who gave birth just before and just after those cutoffs to identify reform effects [Schönberg and Ludsteck, 2014; Kleven *et al.*, 2024b]. In the following, we describe our three sets of results in more detail.

Set (i): Child penalties in Germany. Before turning to earlier decades, as a benchmark, we first report child penalties for mothers delivering their first child after the German reunification (i.e., between 1993 and 2008). The analysis based on monthly data allows us to pinpoint more precisely than previous studies when and how strongly mothers respond to childbirth: Two months before the birth of the first child, just after mandatory maternity leave begins, mothers lose nearly 100% of the potential (counterfactual) earnings they would have without children. The biggest increase after birth appears exactly when the parental leave period ends (36 months after birth), but their earnings never recover to the counterfactual level. The corresponding long-run penalties are substantial: Even ten years after birth, earnings are, on average, about 60% lower than in a counterfactual scenario without children.[4]

Next, we study the evolution of child penalties in West Germany since the 1960s.

[3] Alternatively, some researchers estimate the effects of children using instrumental variable approaches based on (i) the success of in vitro fertilization [Lundborg *et al.*, 2017], (ii) occurrences of twin births [Angrist and Evans, 1998], or (iii) the gender composition of siblings [Angrist and Evans, 1998]. The drawback of these approaches is that they only provide effects for certain compliers and specific samples but not for the entire population. The availability of different methods also fuels an ongoing debate about the optimal strategy to estimate child penalties [Bensnes *et al.*, 2023; Lundborg *et al.*, 2024; Melentyeva and Riedel, 2023]. Some strategies yield smaller effects than the standard approach [Lundborg *et al.*, 2024], and others larger ones [Melentyeva and Riedel, 2023]. Given the lack of methodological consensus, we follow the established mainstream. This strategy also ensures that our results are comparable to the broader literature.

[4] The effect of parenthood varies greatly between both parts of Germany. It amounted to around 64% in the West and 36% in the East.

To that end, we estimate separate event studies for mothers who gave birth in the 1960s, 1970s, and so on. The key insight from this analysis is that the penalties increased over the decades. Mothers who gave birth in more recent decades not only delayed their return to the labor market but also lost a higher percentage of their potential earnings in the long run. For example, in the tenth year after birth, mothers giving birth in the 1960s faced a child penalty of about 35%, while mothers delivering in the 2000s experienced a penalty of approximately 62%. This pattern is general in the sense that we observe it across various subgroups (e.g., within all educational levels).

Set (ii): Decomposition of gender inequality in earnings. The increasing child penalties raise the question of how much of the overall earnings inequality between men and women can be attributed to parenthood. To explore this topic, we first apply the standard approach of Kleven *et al.* [2019b] to our setting and decompose gender inequality in earnings into a *child-related* and a *child-unrelated component*. While overall inequality declined in Germany, child-related inequality increased substantially. For example, women aged 25 to 45 earned about 71% less than men in 1980 and about 54% less in 2013. At the same time, the part of inequality that is related to children went up from about 10% to 34% of men's earnings. As a result, the share of gender inequality we can relate to children increased by a factor of close to five (from about 14% to 64%).

To understand the driving forces behind this development, we next extend the standard decomposition approach. Specifically, we show that child-related inequality grows in three factors: the *child penalties*, *mothers' potential earnings* relative to men (higher-earning mothers have more to lose), and the *proportion of women with children* (a larger share of mothers implies more aggregate losses). In Germany, decreasing fertility rates slightly dampened the evolution of child-related inequality. By contrast, both growing child penalties and rising potential earnings substantially contributed to the increase in inequality. From a normative perspective, only the increase in child penalties appears worrying, while the rise in potential earnings is arguably welcomed. Thus, a crucial insight from our analysis is that without this secular earnings growth, child-related inequality would have increased much less. We conclude that sidelining mothers after childbirth becomes increasingly costly as their earnings potential grows.

Set (iiia): Effects of parental leave reforms. We continue by exploring the factors that cause the rise in child penalties. One natural candidate is family policies, especially parental leave schemes [see, e.g., Olivetti and Petrongolo, 2017].[5] With this in mind, we examine how six parental leave reforms, implemented between 1979 and 1992, impacted child penalties and child-related inequality. The reforms gradually extended job-protected leave from two to 36 months. We find that each of these reforms significantly increased mothers' child penalties in the short run (i.e., during the expanded leave period). The reason is that most mothers responded to the reforms by extending their leave to match the new maximum duration available. However, the reforms did not significantly affect mothers' medium- and long-run earnings penalties. This result aligns well with the findings of Schönberg and Ludsteck [2014]. We then integrate these estimates into our decomposition framework. Specifically, we simulate gender inequality in a counterfactual scenario where job-protected parental leave remained at two months. The novel insight from this analysis is that the short-run effects of all parental leave reforms cumulatively lead to significant shifts in child-related (and overall) gender inequality. According to our estimates, in 2006, child-related inequality would have been about 6.2 percentage points lower without the reforms. In other words, the reforms explain approximately 29% of the rise in child-related inequality between 1980 and 2006. The core logic for how short-run effects accumulate into substantial gender inequality is that the policies impacted most women, as most have children, and removed them from the workforce for a considerable time (36 months instead of two).

Set (iiib): Effects of the 2007 parental benefit reform. In the last step, we demonstrate that policymakers successfully counteracted the trend of increasing child-related inequality with a parental benefit reform. For parents of children born after December 31, 2006, the reform substituted a small means-tested transfer available for up to 24 months after birth with more generous payments limited to 12 months post birth. We find that the reform achieved its stated goals of expediting mothers' reentry into the labor market after childbirth. Indeed, many affected mothers resumed work after twelve months. This prompt return not only facilitated their immediate reintegration but also positively impacted their long-term earnings. For example, in the tenth year after giving birth, mothers' child penalties were approximately 9 percentage points lower than they would have been without the reform. These enduring effects reduced

[5] Theoretically, the impact of parental leave schemes on gender inequality is unclear. On the one hand, job-protected parental leave can strengthen mothers' labor market attachment and increase their labor force participation. These effects could help them maintain or grow their earnings. On the other hand, overly generous leave schemes could reduce mothers' career investments, leading them to stay home longer. This extended time away from work can lower their short-term earnings and cause long-term income losses (e.g., due to diminished human capital).

gender inequality. Our estimates indicate that, in 2018, child-related inequality would have been roughly 1.3 percentage points higher without the 2007 reform. Put differently, child-related inequality would have risen by 13.6% more than it did with the reform.[6]

Related Literature. Our paper contributes to two strands of literature. First, it adds to the extensive literature on gender inequality in earnings [surveyed by Waldfogel, 1998; Olivetti and Petrongolo, 2016; Blau and Kahn, 2017] and, specifically, to recent work highlighting the role of children [Bertrand *et al.*, 2010; Cortés and Pan, 2023; Kleven *et al.*, 2019b,a; Kuziemko *et al.*, 2018]. The seminal papers by Angelov *et al.* [2016] and Kleven *et al.* [2019b] establish an event-study approach to estimate child penalties. Many later papers use their approach to investigate child penalties in various countries [see Cortés and Pan, 2023, for a survey]. One example is Kleven *et al.* [2019a], who estimate child penalties across countries using an integrated framework. Most relevant to our paper, they document (based on survey data) that the current earnings penalties are larger in West Germany than in the UK, US, Sweden, and Denmark. While most existing papers focus on child penalties in recent decades, a few exceptions exist. Kleven *et al.* [2024b], for example, examine the evolution of child penalties in Austria. Most recently, Kleven *et al.* [2024a] use pseudo-event studies based on cross-sectional (survey or census) data to estimate child penalties in employment for more than 100 countries. In some cases, they are even able to analyze data spanning multiple centuries. We add to this literature by documenting how child penalties in Germany evolved since the 1960s (using monthly administrative data) and by examining how children contributed to overall inequality. Moreover, we introduce an extended decomposition approach that breaks down child-related earnings inequality into the aforementioned components. Our approach highlights the role of potential earnings and provides a better understanding of the forces driving gender inequality.

Second, we contribute to a literature that discusses the effects of family policy on (female) labor market outcomes [summarized by Olivetti and Petrongolo, 2017; Canaan *et al.*, 2022]. For example, Schönberg and Ludsteck [2014] study the effects of German parental leave reforms on maternal labor supply.[7] A much smaller set of

[6] Note that we only observe mothers for 12 years after the reform's implementation. The longer-run effects on overall gender inequality are probably larger. One reason is that the positive long-run effects on earnings may continue to unfold into later years after childbirth. Additionally, as time progresses, the data encompasses (i) more cohorts impacted by the reform and (ii) an expanded group of mothers we observe for longer periods after birth. Both forces amplify the observed effects. This feature may also explain why the leave reforms have a greater measured impact on inequality than the benefit reform.

[7] Bauernschuster and Schlotter [2015], Busse and Gathmann [2020], and Gathmann and Sass [2018]

papers focuses on the effects of parental leave reforms on gender inequality or child penalties. Kleven *et al.* [2024b] investigate the impact of several parental leave reforms since 1990 in Austria; Andresen and Nix [2023] examine the effects of paternity leave reforms in Norway; and Bailey *et al.* [2024] study California's 2004 Paid Family Leave Act. All these papers find negligible effects of parental leave policies on child penalties or child-related inequality.[8] Our contribution is to evaluate the impacts of two types of policies in a single setting: parental leave and parental benefit schemes. Similar to previous studies, we find that the German parental leave reforms only boosted child penalties in the short run. However, in our setting, these short-run effects translate into non-negligible impacts on child-related inequality. A potential reason why the German parental leave policies have larger effects is that they stand out in magnitude: The reforms extended parental leave from two to 36 months. By contrast, the 2007 benefit reform introduced more generous income replacements but limited them to a shorter period after birth. In line with the idea that such monetary incentives shortly after childbirth matter [Gruber *et al.*, 2023; Kuka and Shenhav, 2024], the reform successfully reduced child penalties both in the short and the long run. To our knowledge, our paper is the first to demonstrate that certain family leave policies can have long-lasting effects on maternal earnings.[9]

Outline. The remainder of our paper evolves as follows. Section 2 provides an overview of the institutional background and the data. Section 3 presents our estimates of child penalties in Germany. Section 4 introduces and applies our decomposition approach. Section 5 analyzes the impacts of the six parental leave reforms, and Section 6 focuses on the 2007 parental benefit reform. Finally, Section 7 concludes.

2 Institutional Background and Data

This section describes the institutional background and the data.

study the impacts of German childcare policies on maternal labor supply.

[8] A handful of papers also examine the impact of formal childcare on child penalties. Kleven *et al.* [2024b] focus on Austria, Krapf *et al.* [2020] on Switzerland, and Huber and Rolvering [2023] on Germany. The results vary across settings, potentially depending on how strongly parents substitute informal childcare for formal options.

[9] Instead, studies have shown that other policies can increase child penalties in the long run. For example, Gruber *et al.* [2023] find that a Finish policy paying mothers to stay home with children under three reduces maternal employment in both the short and long term. Like the German parental benefit reform, the Finish policy significantly altered monetary incentives for mothers.

2.1 Institutional background

Labor market. Two features of the German setting make it an interesting case for studying gender inequality. First, the German labor market demonstrates substantial gender disparities. For example, in 2018, the unadjusted gender wage gap was 20.1%, making it one of the largest in Europe [Mischler, 2021]. Nonetheless, women's employment rate in the same year stood at a relatively robust 72.1%, compared to 75.9% for men. A striking difference also emerges in part-time employment rates: 47.9% of employed women worked part-time in 2018, compared to only 11.2% of men. Second, as in other countries, Germany has witnessed a significant decline in gender gaps over time. It seems natural to examine the forces behind the changes in gender inequality.

Gender norms. Historically, Germany was a rather gender-conservative society with strong gender identity norms [Akerlof and Kranton, 2000]. While gender norms in Germany apparently became more progressive over time (see Appendix Figure A.1), it is still on the more conservative side in international comparison (see Appendix Figure A.2). These patterns are, for example, observable in ISSP survey data reflecting opinions on family and gender roles.[10] Consequently, our study offers insights into the role of parenthood and family policies in an environment where traditional gender norms were and still are more deeply ingrained than in other countries [Kleven *et al.*, 2019a].

East vs. West Germany. There is a remarkable within-country heterogeneity. West Germany traditionally tended to support a male-breadwinner model, with low female labor force participation and low fertility rates. By contrast, socialist East Germany propagated a dual-earner model. Here, mothers typically worked full-time. In line with these differences, gender norms also differed strongly between East and West Germany [Campa and Serafinelli, 2019; Becker *et al.*, 2020]. Until today, many of those differences seemingly remained [Kreyenfeld and Geisler, 2006; Jessen, 2022]. For example, in 2002, only 14.5% of West German mothers with children between three and six worked full-time [Kreyenfeld and Geisler, 2006]. In East Germany, this rate was 50.5% in the same year. Moreover, norms show some persistence. Specifically, East German gender norms seem to last over time, while some West Germans tend to adopt Eastern standards regarding labor market behavior after childbirth [Boelmann *et al.*, 2024]. Appendix Figure A.1 presents additional suggestive evidence in line with the

[10] The International Social Survey Programme (ISSP) represents a collaborative effort among various countries to conduct surveys on topics pertinent to social science research.

idea that, even after the reunification, norms differed in the East and the West.

Family policies. Mothers' supply of work in the labor market likely depends on how strongly family policies do or do not "support" them. Arguably, West German institutions did not facilitate maternal labor supply in the workforce for a long time, while East German institutions aimed at promoting a more equal division of work in the labor market.[11] One example pointing in this direction is the parental leave regulations. As we will discuss in detail, between 1979 and 1992, a series of West German paid parental leave reforms expanded the job-protected leave duration from 2 to 36 months post birth [Schönberg and Ludsteck, 2014]. In East Germany, instead, the option to take parental leave was limited to at most one year in most cases. A similar picture arises regarding the availability of formal child care [Bauernschuster and Schlotter, 2015]. In line with the idea that mothers are the primary caregivers, both the availability of formal childcare and enrollment rates were low in West Germany for a long time. This holds especially for children under the age of three. For them, the average childcare coverage only started to increase in the early 2000*s* – from 1.7% in 1998 to 29.3% in 2019 [Huber and Rolvering, 2023]. In the East, instead, the provision of public care for children between one and six years was higher in the former GDR, and remained so after the reunification. For example, in 2002, about 35% of children below the age of three had access to public child care in East Germany.

2.2 Data

Our analysis utilizes administrative pension-insurance data provided by the *German Pension Insurance* (Deutsche Rentenversicherung, DRV). The DRV collects detailed longitudinal data in spell form, holding all the information needed to calculate pension claims for the majority of Germans (approx. 85% of employed persons). This detailed data collection stems from the requirement that the vast majority of German employees are mandatorily insured in a national pay-as-you-go pension scheme of a Bismarckian type.[12]

Sample and content. For our project, we obtained a random sample from the population of people insured with the DRV (called *Versicherungskontenstichprobe* or VSKT).

[11] A more balanced division in the labor market does not necessarily lead to a fairer distribution of housework. In fact, East German women also perform more housework than East German men [Jessen *et al.*, 2024].

[12] There is a separate pension system for civil servants. Thus, neither this group nor self-employed people appear in the pension register data.

Specifically, the VSKT is based on a 0.5% stratified random sample of the insurees born in each calendar year since 1935. Among other things, the data contain complete *monthly* earnings biographies, unemployment spells, or periods of child care. It also provides individual characteristics (such as the insured person's birthdate and the birthdates of their children). Other potentially useful characteristics, including a panel of zip code information or the place of birth, are not part of the data. To qualify for inclusion in the data, insurees must (i) live in Germany, (ii) have at least one entry in their pension insurance records, and (iii) be aged between fourteen and sixty-seven. Bönke *et al.* [2015] provide a more detailed description of the pension insurance data (see also Supplemental Materials).

Data merge and sample restrictions. In our analysis, we merge VSKT data from the years 2002 and $2004 - 2018$.[13] Note that, for each annual wave of the VSKT, the DRV draws an independent 25% subsample of the random sample mentioned above. Consequently, some people might be part of multiple waves, but there are no time-consistent individual identifiers. To construct a consistent dataset, we employ a record-linkage approach that matches individuals on their monthly employment biographies to eliminate duplicate observations. Our final sample contains all monthly entries from 1949 to 2018 for all cohorts born in 1935 or later. Because there are structural breaks in the data before 1960, we focus on the period 1960 to 2018.[14] Moreover, we restrict our analysis to individuals whose pension records have been verified and completed by the DRV.[15]

Benefits of the pension insurance data. Three aspects of our data merit special attention. First, unlike other German datasets, such as those from the German Institute for Employment Research, we can precisely track if and when mothers have children. The pension insurance collects these data because mothers' pension claims depend on children. Instead, the data only allow us to match children to fathers for a small, selected subset. Therefore, our primary VSKT-based analyses focus on mothers.

[13] The research center of the pension insurance did not release a VSKT wave for 2003.

[14] Most importantly, the German government implemented a structural pension insurance reform in 1957. This reform fundamentally changed the nature of the data. There are two further data limitations. First, the data does not contain one-time payments before 1984. Second, the data is top-coded, as employees contribute a share of their gross wage to the pension insurance up to a ceiling. Only very few women have earnings above this ceiling, though. In robustness checks, we impute one-time payments and earnings above the earnings ceiling following Bönke *et al.* [2015]. The results are unchanged.

[15] The DRV conducts a verification process for all insurees at the age of 30. The goal is to ensure all relevant data for future pension calculations are captured. We focus on these individuals in our analysis because their insurance histories are more reliable.

To assess paternal impacts, we utilize the *German Taxpayer Panel* in supplementary analyses. Second, our dataset also encompasses earnings histories from individuals in East Germany before reunification. The reason is that the GDR implemented a similar pension system, and after reunification, the (West) German pension insurance incorporated East German data to calculate pension entitlements. However, when we analyze the pre-unification periods, our primary focus remains on West Germany. The reasons are the complexities and potential inaccuracies in interpreting earnings from the GDR. Third, we observe earnings on a monthly level. This feature enables us to track labor-market trajectories in a granular way.

3 Child penalties in Germany

We explain the method for estimating child penalties in Section 3.1. Section 3.2 follows with our results on child penalties for West- and East Germany in the recent decade. Finally, in Section 3.3, we track how these penalties changed from the 1960s to the 2000*s*.

3.1 Event study methodology

The first-best strategy to investigate the effect of childbirth on mothers' earnings trajectories would be to use random assignment of childbirth. Naturally, such experiments are infeasible. Recognizing this limitation, Angelov *et al.* [2016] and Kleven *et al.* [2019b] propose an event-study approach that nevertheless allows researchers to estimate the effects of childbirth, which we adapt in this paper. The central idea of this approach is that childbirth introduces abrupt drops in mothers' earnings. Even though mothers do not make fertility decisions randomly, these sudden changes are arguably orthogonal to other unobserved earnings components. The approach then exploits these changes surrounding the birth of the *first* child to trace out the overall dynamic impacts of childbirth on earnings and inequality. A major advantage of the method is that it allows us to estimate effects for the full population.

Estimating event studies around childbirth. The details of the estimation approach are as follows. We focus on the sample of mothers m and define s as the calendar month and b_i as the birth month of Mother i's first child. Consequently, for mother i, $t = s - b_i$ is the time relative to the birth of the first child, measured in months (called event time). The running indices s and t increment sequentially without resetting at the start of each year. Using data around childbirth, we estimate the following regression with ordinary least squares, where Y_{is}^{m} denotes mother i's earnings in month s:

$$Y_{is}^{m}=\sum_{t\neq-12}\alpha_t\cdot\mathbb{1}[t=s-b_i]+\sum_k\beta_k\cdot\mathbb{1}[k=age_{is}]+\sum_y\gamma_y\cdot\mathbb{1}[y=year_s]+u_{is}. \quad (1)$$

The first term in equation (1) captures the impacts of event time dummies; we omit the event time dummy at $t=-12$ as a reference period. The second term captures life-cycle effects with monthly age dummies, where age_{is} is mother i's age in month s. The third term controls flexibly for macroeconomic effects with year dummies: $year_s$ is the calendar year associated with month s. Notably, the estimated coefficients $\widehat{\alpha}_t$ identify the total impact of childbirth on earnings at event time t under one key assumption: The non-child earnings path must be smooth conditional on controls.

We then define the *child penalty* at event time t as:

$$P_t=\widehat{\alpha}_t/E[\tilde{Y}_{is}^{m}|t], \quad (2)$$

where $\tilde{Y}_{is}^{m}$ is an estimate of the counterfactual earnings of mother m in the absence of children, and the expectation operator indicates that we take an average over all mothers that are at event time t. Equation (2), hence, translates the absolute loss in earnings $\widehat{\alpha}_t$ into a percentage loss relative to mothers' counterfactual earnings. Formally, these counterfactual earnings are:

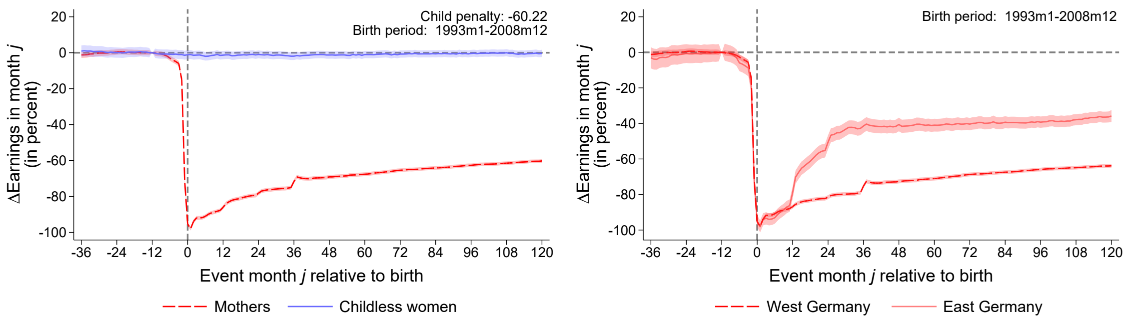
$$\tilde{Y}_{is}^{m}=\sum_k\widehat{\beta}_k\cdot\mathbb{1}[k=age_{is}]+\sum_y\widehat{\gamma}_y\cdot\mathbb{1}[y=year_s]. \quad (3)$$

Three further points are of note. First, although model (1) focuses on the event of having the first child, the longer-run penalties P_t encompass the effects of subsequent children. Second, we apply population weights in our analysis to correct for the oversampling of certain subgroups in the German pension data. Hereby, we ensure our results are representative of the full population. Third, we follow Kleven *et al.* [2019b] and estimate (1) on a balanced sample of mothers, whom we constantly observe at least three years prior to and ten years after the birth of the first child.

Childless women as comparison group. The event study methodology aims to identify the causal effects of children on maternal earnings. A potential threat to identification is that the event time dummies could pick up general trends in female earnings unrelated to childbirth that our age and year dummies fail to capture. We respond to this concern by using women without children as a comparison group. To

Figure 1: Child penalties for births between 1993 and 2008

(a) Entire Germany **(b) West Germany versus East Germany**



Notes: This figure presents the estimated child penalties in mothers' earnings, $\widehat{P}_t$. In Figure A.3a, the dashed red line shows the child penalties for the entire country of Germany (including the Western and Eastern regions). The solid blue line represents the corresponding placebo effects for childless women. In the upper right corner of each figure, we report the child penalties at event time 120 (i.e., ten years after birth). Figure A.3b depicts child penalties in East and West Germany separately. Both figures cover first births between 1993 and 2008 and consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

achieve close comparability between mothers and childless women, we use a matching strategy that assigns pseudo-birth events to childless women [Kleven *et al.*, 2019b]. Appendix A describes this approach in detail. As will get clear in Section 3.2, both groups' earnings evolve in parallel before the (pseudo) event. Moreover, the estimated point estimates of the "pseudo child penalties" are statistically indistinguishable from zero for all event times (and statistically insignificant). These findings speak against the just-stated concerns.

3.2 Baseline estimates of child penalties

This section estimates child penalties for German mothers in the post-reunification period as a benchmark. Specifically, we focus on women with first births (or placebo births) between 1993 and 2008 and trace their earnings for ten years after birth (which is impossible for births after 2008).

Main results. Figure A.3a depicts our baseline estimates of the child penalties in earnings for German mothers. The dashed red line represents the percentage child penalty P_t for each event time from 36 months before to 120 months after birth. In the top right corner, we also report the penalty in $t = 120$ (i.e., ten years after birth). The solid blue line focuses on childless women and shows the penalties for pseudo-births. Figure A.3b, instead, depicts child penalties separately for West (dashed red line) and East German mothers (solid red line).[16]

Three main messages emerge. First, the child penalties for German mothers are substantial (see Figure A.3a). Most mothers lose all of their earnings around childbirth. Importantly, this decline in income is not only a temporary setback. Mothers' earnings fail to recover fully and, instead, stabilize at around 40% of its counterfactual level without children. Ten years (120 months) after giving birth, German mothers still experience a child penalty in earnings of about 60% (see Figure A.3a).

Second, the figure visualizes the advantage of having monthly data: We can precisely pinpoint when mothers' earnings drop or increase. Labor earnings fall sharply two months before birth (right after the start of the maternity protection period) and rise most significantly 36 months after birth (when the parental leave period ends. This pattern suggests that parental leave policies are pivotal in shaping mothers' labor supply decisions, a finding we more formally confirm in Section 5.

Third, as also documented by Jessen [2022], mothers in West and East Germany behave very differently, even those who gave birth after the reunification in 1990. In $t = 120$, for example, East German mothers face much smaller child penalties in earnings (36%) compared to their West German counterparts (64%). Various factors may explain these regional differences between the East and West, and the literature already discussed some of these differences. One potential explanation is that the culture in which a woman grows up influences her labor market decisions later in life [Boelmann *et al.*, 2024]. Alternatively, the divergence might stem from the incomplete convergence of East and West German institutions. For example, the availability of childcare has historically been much greater in the East than the West [Gathmann and Sass, 2018; Busse and Gathmann, 2020]. Such differences may explain the differential ability of mothers in the East and the West to participate in the labor market.

Further analyses. We provide additional analyses in the Appendix. First, we demonstrate that child penalties vary across certain groups. For example, they correlate with educational attainment (see Appendix Figure A.4). Mothers lacking formal training encounter the largest penalties. Those with vocational training or a high school diploma experience slightly less pronounced adverse impacts, and the penalties are the smallest for mothers holding a university degree. Subsequent fertility also plays a critical role. Appendix Figure A.5 splits the sample by the total number of children

[16] Appendix Figure A.3 presents similar figures studying the extensive margin (labor market participation). The figures are based on regressions that use a dummy indicating whether a mother has strictly positive earnings as an outcome. The penalties in this dimension are also substantial.

(one, two, and three or more). In line with the notion that mothers' focus on parenthood intensifies when they have more children, we find that long-run earnings drop more after the first child for mothers who go on to have more children. Note that, as always, these and other sample splits in our paper do not necessarily offer a causal interpretation. The splitting variable may correlate with unobserved factors that cause the variation in child penalties.

Second, given that our focus has been on women until now, one may wonder how parenthood impacts German men. To study men, we use yearly tax data, the German Taxpayer Panel (see Appendix Section B for a description of this dataset). Applying the same method to these data, we find that childbirth does not impact fathers' earnings (see Appendix Figure A.6). This finding aligns with the prevailing consensus in the literature.

Third, we examine how the child penalties evolve in the very long run (i.e., over 20 years after the birth of the first child). This topic is interesting because, as children grow older and become more independent, they typically require fewer parental inputs. Mothers might then face fewer constraints on their labor supply. However, speaking against this idea, women continue to suffer from substantial earnings losses even two decades after childbirth (see Appendix Figure A.7). We need to take these results with a grain of salt: Given the long period from birth to the point of outcome evaluation, the identifying (smoothness) assumption seems stronger.

3.3 Historical Perspective

Several studies highlight substantial child-related earnings gaps in recent decades across various countries [Angelov *et al.*, 2016; Kleven *et al.*, 2019a,b; Andresen and Nix, 2022, 2023]. Yet, our understanding of how these disparities have evolved over time remains limited. This subsection explores if the child penalties in Germany have always been as pronounced as they are today, leveraging our pension data starting in the 1960s.[17] We concentrate all subsequent analyses, including the decomposition analysis and the reform evaluations, on West Germany. This is because the study period covers the era of the former GDR, where there is more uncertainty about the data quality.

Main results. Figure 2 depicts the evolution of the child penalties from the 1960*s* to the 2000*s*. For this analysis, we segment the data by the children's birth decade and

[17] Technically, the data even starts in the 1950s. However, as already discussed, multiple data breaks in the 1950s impede a clean analysis.

estimate the regression equation (1) separately for each decade. Figure 2a shows results for mothers giving birth to their first children in the 1960s, Figure 2b for mothers with children in the 1980s, and Figure 2c for mothers delivering in the 2000s.[18] Again, the dashed red lines (solid blue lines) show the estimated (placebo) penalties. The figures representing the 1980s and 2000s also include the penalties from the 1960*s* for comparison (solid light red lines).

Several findings emerge. First, the patterns in the 1960s are very different from those observed in recent decades: While the earnings of women who gave birth in the 1960*s* also experienced a substantial decline around the time of childbirth, they quickly rebounded two months after birth. Over the long term, there is also more convergence towards previously observed earnings levels. Ten years after birth, for example, mothers' earnings were only about 35% lower than their counterfactual earnings without children. Second, the picture changed drastically in the 1980s. In the first six months after birth, mothers faced much larger child penalties (close to 100%). However, compared to the 1960s, there was also a greater increase in the long-run child penalties. The child penalty amounted to roughly 56% in the tenth year after birth. Third, the results for the 2000s are even more extreme. We find child penalties to be much more considerable in the short, medium, and long run. Ten years after birth, for example, the child penalty was 62%.[19]

Role of education. Appendix Figures A.14-A.16 show that, even within all educational levels, the child penalties have consistently risen between the 1960s and the 2000s. We can interpret this result in two ways. First, it suggests that the growth in child penalties is a general pattern we observe across all educational levels. Second, the heterogeneity indicates that, although women's education levels have continuously increased since the 1960s, this trend alone cannot explain the rise in child penalties.

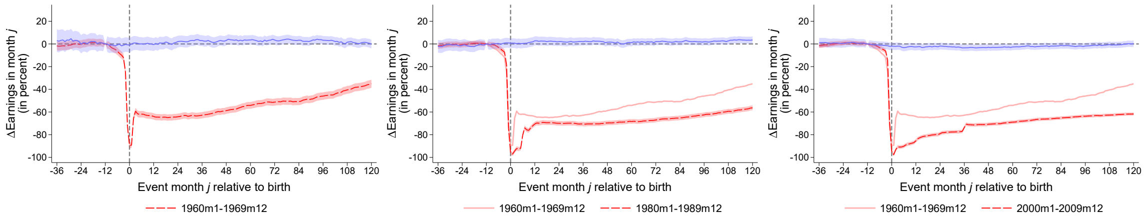
Next steps. Two questions emerge naturally from our results that motivate our next steps. First, the drastic increase in child penalties over recent decades prompts the question of its implications for overall gender inequality. Does this notable increase imply that childbirth now contributes more significantly to the overall gender inequal-

[18] In the Appendix, we provide figures for all decades (including the 1970 and 1990s). Specifically, Figure A.8 shows the child penalties in earnings and Figure A.9 the extensive margin child penalties. Moreover, Figures A.10 and A.11 demonstrate the very long run penalties (20 years) for all decades. Lastly, Figures A.12 and A.13 report results for East Germany. In the 1960s, the penalties were relatively comparable in the West and the East. However, they diverged afterward and never converged again. As noted earlier, we do not explore these trends further due to potential data quality issues.

[19] This development in Germany differs from that in other countries. For example, the US [Kleven, 2023], child penalties decreased in the last decades.

Figure 2: Child penalties in West Germany in different decades

(a) The 1960s (b) The 1980s (c) The 2000s



Notes: The figures show the estimated child penalties in the earnings of West German mothers giving birth in the 1960s (Figure 2a), the 1980s (Figure 2b), and the 2000s (Figure 2c). In each figure, the dashed red line depicts child penalties for actual mothers, while the solid blue line shows placebo child penalties for pseudo births of childless women. The shaded areas depict 95% confidence intervals. In Figures 2b and 2c, the solid light red lines display the child penalties from the 1960s for comparison.

ity we observe today? Section 4 explores this topic in detail. Second, the increase also raises the question of what forces have caused the increase in child penalties. A quick inspection of Figures 2a-2c offers a clue: All these figures reveal notable upward jumps in earnings at certain points in time, indicating that many mothers simultaneously re-enter the workforce. In fact, over the years, the German government has implemented a series of reforms extending parental leave, and these discontinuities align with the maximum duration of parental leave applicable at the respective times. Section 5, therefore, investigates the role of parental leave and parental benefit reforms in detail.

4 Decomposition of gender inequality

This section investigates to what extent parenthood contributed to the earnings inequality between German men and women and how this contribution changed over time. For this purpose, we follow Kleven *et al.* [2019b] and decompose the gender earnings gap into a child-related gap and child-unrelated inequality. We then extend the approach and further decompose child-related inequality.[20]

4.1 Decomposition approach

Standard decomposition of overall gender inequality. For a given population, *overall gender inequality in earnings* Δ_y^o (i.e., the raw gender gap) is commonly measured as the difference between men's and women's average earnings, divided by

[20] See Appendix C.1 for further details.

men's average earnings. Kleven *et al.* [2019b] decompose Δ_y^o into one part attributable to children (*child-related inequality*, Δ_y^r) and one child-unrelated part (*child-unrelated inequality*, Δ_y^u):[21]

$$\begin{aligned}
\Delta_y^o &= \frac{E[Y_{is}^{men}|y] - E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100 \\
&= \underbrace{\frac{E[\tilde{Y}_{is}^{women}|y] - E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100}_{\Delta_y^r} + \underbrace{\frac{E[Y_{is}^{men}|y] - E[\tilde{Y}_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100}_{\Delta_y^u} . \quad (4)
\end{aligned}$$

In equation (4), $E[\tilde{Y}_{is}^{women}|y]$ denotes the average potential earnings of women in the counterfactual state of the world without children. Note that, for childless women, these counterfactual earnings simply equal their observed earnings. For mothers, by contrast, counterfactual and observed earnings differ due to the effects of children (i.e., due to child penalties).[22]

Estimating counterfactual earnings. As apparent in (4), the decomposition requires an estimate of mothers' potential (counterfactual) earnings of mothers. In principle, as explained in Section 1, one can get such an estimate by estimating model (1). However, in the following, our goal is to allow the counterfactual estimates to vary over time. To that end, we estimate regression (1) separately for each birth cohort (i.e., for each set of mothers giving birth to their first child in a given calendar year $g_i \in [1954, 1955, ..., 2018]$).[23] We, hence, model the relation between the observed and potential earnings at event time t for a mother i giving birth in year g^i as:

$$Y_{is}^m = \tilde{Y}_{is}^m + \widehat{\alpha}_t^{g_i} + \hat{u}_{is}, \quad (5)$$

where $\widehat{\alpha}_t^{g_i}$ captures the estimated effect of the first child on maternal earnings in t.[24]

[21] As noted by Kleven *et al.* [2019b], this measure of child-related inequality only captures the effects of children that realize around and after childbirth (i.e., responses to the realization of motherhood). Potential anticipatory effects of parenthood that are unrelated to the timing of birth (e.g., education decisions related to the future fertility plans) are included in Δ_y^u.

[22] An underlying assumption is that fathers' earnings are not affected by children. This assumption is consistent with our results from the German Taxpayer Panel (see Appendix Figure A.6).

[23] We can run cohort-specific regressions, including calendar year dummies, due to the availability of monthly observations. With yearly earnings information only, event time and calendar year would be perfectly collinear within each birth cohort.

[24] Mother i's predicted earnings absent kids in year y are $\tilde{Y}_{is}^m + \widehat{u}_{is} = \widehat{\beta}_k^g + \widehat{\gamma}_y^g + \widehat{u}_{is}$ if she is k years old.

Extended decomposition approach. Next, we extend the standard decomposition approach of Kleven *et al.* [2019b]. Our goal is to rewrite Δ_y^r such to obtain more general insights into the drivers of child-related inequality. We start from the observation that child-related inequality only depends on the effects of children on mothers' earnings. Building on this insight and denoting the share of mothers in a given year by ϕ_y, we can use (5) to express Δ_y^r as:

$$\begin{aligned} \Delta_y^r &= \frac{E[\tilde{Y}_{is}^{women}|y] - E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100 \\ &= \phi_y \cdot \frac{E[\tilde{Y}_{is}^{m} + \widehat{u}_{is}|y] - E[Y_{is}^{m}|y]}{E[Y_{is}^{men}|y]} \times 100 \\ &= -\phi_y \cdot \frac{E[\widehat{\alpha}_t^{g_i}|y]}{E[Y_{is}^{men}|y]} \times 100. \end{aligned} \tag{6}$$

We can then rearrange (6) to get an expression with a straightforward interpretation. For that, we expand it by the average counterfactual earnings of mothers, $E[\tilde{Y}_{is}^{m}|y]$, and define two terms: (i) the average child penalty across all mothers in year y,[25]

$$\bar{P}_y = \frac{E[\widehat{\alpha}_{is}|y]}{E[\tilde{Y}_{is}^{m}|y]}, \tag{7}$$

and (ii) the ratio of mothers' counterfactual earnings to men's earnings in year y,

$$\Psi_y = \frac{E[\tilde{Y}_{is}^{m}|y]}{E[Y_{is}^{men}|y]}. \tag{8}$$

Substituting both terms into (6), we end up with the simple expression:

$$\Delta_y^r = -\phi_y \cdot \bar{P}_y \cdot \Psi_y \times 100. \tag{9}$$

Equation (9) clarifies that child-related gender inequality depends on three factors: the share of mothers ϕ_y, their average child penalty $\bar{P}_y$, and the counterfactual earnings ratio Ψ_y (a relative measure of mothers' earnings potential). We can investigate the development of these factors over time to illustrate the sources of variations in child-related inequality.

This decomposition of child-related inequality is helpful for several reasons. First, it not only verifies that child penalties affect child-related inequality but also demonstrates precisely how this impact unfolds. Second, perhaps more interestingly, the

[25]Note that the average child penalty $\bar{P}_y$ depends on the event-specific child penalties and on the distribution of (pre-birth and post-birth) event times across all mothers in year y.

decomposition also reveals that, even if child penalties remain constant, child-related inequality can change over time. In particular, if the earnings potential of mothers Ψ_y grows, then child-related inequality goes up ceteris paribus. Intuitively, in this case, women face larger potential earnings losses from child-related career breaks. Note that variations in ϕ_y, $\bar{P}_y$, and Ψ_y have quite different normative implications: While growing child penalties are arguably worrying, an increase in the counterfactual earnings of mothers implies an improvement in gender equality. An increase or decrease in the fertility rate is less directly related to aspects of gender inequality but may come with other challenges beyond the scope of this paper.

4.2 Sample

We select the sample for our decomposition analyses with two goals in mind. One goal is to ensure that we cover individuals of the same age across all years of the analysis. However, without additional sample restrictions, the age of individuals in our decomposition sample would systematically vary between the initial and later years. The reason is that, due to the construction of our pension register data, earnings information for older (younger) women are missing in earlier (later) years.[26] To ensure that the potential age-composition effect cannot drive our results, we apply two key sample restrictions: We focus our decomposition analysis on the years 1980 to 2013, and we restrict our sample to individuals aged 25 to 45.[27] As a result, the sample includes individuals across the entire age span (25 to 45) in all years (1980-2013). This restriction also implies that we decompose inequality for individuals within an age where having small children is common. Appendix Figure A.20 reports similar results for alternative sample definitions.[28]

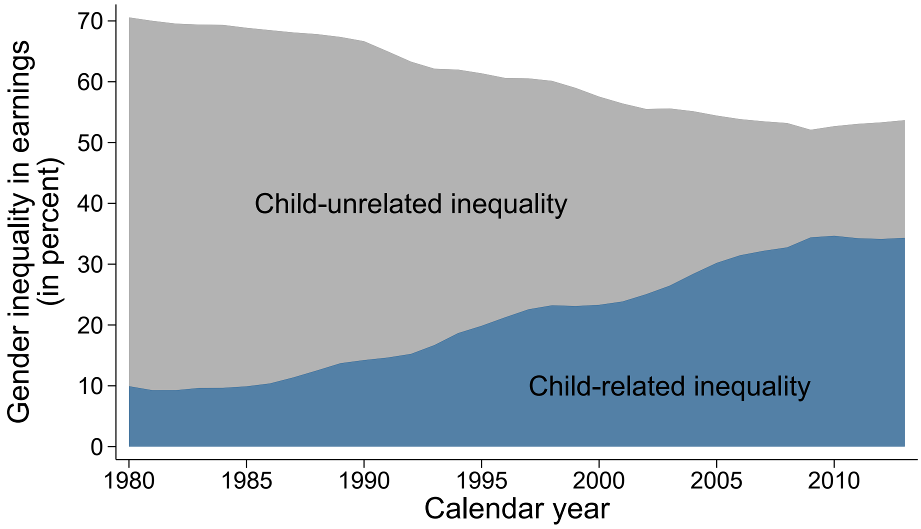
Our second goal is to consider the effects of childbirth for *all German parents* in our decomposition analysis. Our previous child-penalty estimates, however, focused on a restricted (balanced) sample: mothers whom we continuously observe in three pre-birth and ten post-birth years. Thus, our baseline estimates do neither capture long-term effects beyond the tenth post-birth year nor effects for mothers with gaps in their observation periods. To ensure our analysis does not underestimate children's contribution to inequality by neglecting those potentially relevant effects, we follow

[26]Our pension register data covers cohorts born between 1935 and 1988. Thus, all people in the data were aged 25 or younger in 1960. Conversely, all people in the data were 30 or older in 2018 (our latest wave).

[27]One alternative to these sample restrictions is extrapolating certain estimates (e.g., child penalties) to earlier/later years. This strategy would require strong assumptions, which we are unwilling to make.

[28]Specifically, we also perform decomposition analyses for (i) a sample of individuals aged 20 to 40 in the years $1975-2008$ and (ii) a sample of individuals aged 30 to 50 in the years $1985-2018$.

Figure 3: Decomposition of overall gender inequality in earnings



Notes: This figure decomposes the overall gender inequality in earnings into a child-related part (blue area) and a child-unrelated residual part (gray area). Equation (4) formalizes this decomposition. The underlying event study models allow for year-specific event coefficients and control for year and age dummies. The sample spans the years from 1980 to 2013 and includes West German men and women between the ages of 25 and 45.

the approach of Kleven *et al.* [2019a] and re-estimate event studies (i) based on an unbalanced sample of all mothers and (ii) consider event periods beyond $t = 120$.

4.3 Results

In this section, we report the results from the standard decomposition approach of Kleven *et al.* [2019a] and we apply our extended approach to the data.

Results from the standard approach. Figure 3 presents the results of the standard decomposition analysis. It decomposes overall inequality into child-related (blue) and child-unrelated parts (gray) and shows how they evolved. Two key patterns emerge from the data. First, in Germany, overall gender inequality in earnings decreased from about 71% in 1980 to 54% in 2013. Despite this substantial reduction, gender inequality in Germany is still large in international comparison [see, e.g., Kleven *et al.*, 2019a, 2024b]. The second message is, however, that while overall and child-unrelated inequality fell, child-related inequality heavily increased. In 1980, it only amounted to 9.9% (i.e., 14% of overall gender inequality). By contrast, in 2013, it reached 34% of men's earnings (i.e., 64% of overall inequality). Put differently, without children, overall gender inequality would have dropped from 61% to 19%. Our results imply that, in Germany, persisting gender inequalities predominantly stem from children.[29]

Results from the extended approach. In the next step, we investigate why child-related inequality increased. To that end, we apply the extended decomposition approach introduced in Subsection 4.1. According to equation (9), child-related inequality increases in (i) the share of mothers ϕ_y, (ii) the average child penalty $\bar{P}_y$, and (iii) the relative earnings potential of mothers Ψ_y.

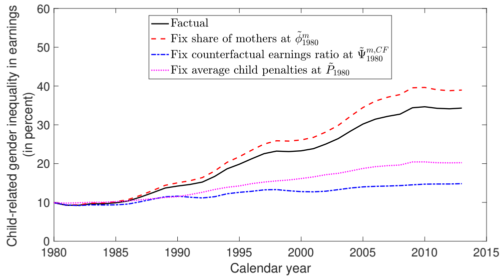
In Germany, all three factors changed substantially between 1980 and 2013 (see Appendix Figure A.17). Specifically, the share of mothers fell from 87% to 77%. Meanwhile, child penalties grew substantially, with the average penalty $\bar{P}_y$ increasing from 32% of men's earnings in 1980 to 55% in 2013. The most dramatic change, however, occurred in the relative earnings potential of mothers: Their counterfactual earnings, as a percentage of men's earnings, more than doubled from 35% to 82% over 33 years. This development is likely the result of improved educational levels among mothers (documented in Appendix Figure A.18) and a general increase in female labor market participation (shown in Appendix Figure A.19). We conclude that while decreasing fertility rates reduced child-related inequality, increases in the other two factors amplified it. However, Appendix Figure A.17 alone does allow not us to determine the precise contributions of these factors to inequality.

Thus, our next step is to more accurately quantify how each factor influences child-related and overall inequality. Specifically, we compute how inequality would have evolved if one of the components (i)-(iii) remained at its value in 1980. Figure 4a illustrates this thought experiment for *child-related inequality*. The solid black line depicts the factual evolution of child-related inequality over time. As we already noted, between 1980 and 2013, it increased by 24 percentage points. Our analysis reveals that if the average child penalty had remained at its 1980 level, instead, child-related inequality would have grown by 10 percentage points, much less than observed (see dotted pink line). Similarly, with a constant earnings ratio Ψ_{1980}, the increase would have been just 5 percentage points (see dash-dotted blue line). Finally, child-related inequality would have grown by 30 percentage points if the share of mothers had remained fixed (see dashed red line). We conclude that the declining fertility rate slightly dampened child-related inequality. By contrast, both growing child penalties and mothers' increasing earnings potential strongly contributed to the increase in

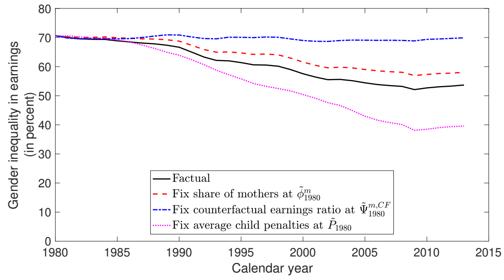
[29] Our findings differ somewhat from previous studies: In Denmark and Austria, the share of overall gender inequality attributable to children also increased over time. In both cases, however, this development was driven by reductions in overall inequality. Child-related inequality as a percentage of average men's earnings remained almost constant [Kleven *et al.*, 2019b, 2024b].

Figure 4: Drivers of child-related inequality in earnings

(a) Child-related inequality



(b) Overall inequality



Notes: Building on our extended decomposition approach, this figure examines the drivers of child-related (Figure 4a) and overall inequality (Figure 4b). It focuses on the period 1980 to 2013. The solid black lines show the factual evolution. The other lines demonstrate the evolution of child-related or overall inequality under the assumptions of a fixed share of mothers (dashed red lines), a fixed average child penalty (dotted pink lines), and a fixed ratio of mothers' counterfactual earnings relative to men's (dash-dotted blue lines).

child-related inequality. A crucial insight from our analysis is, thus, that without the growth in mothers' potential earnings, child-related inequality would have increased much less. In other words, sidelining mothers after childbirth becomes increasingly costly as their earnings potential grows.

Figure 4b presents the same thought experiment for *overall inequality*, Δ_y^o: Again, we hold one of the three factors constant, while allowing the others to change. For comparison, note that overall inequality factually decreased by 17 percentage points (see solid black line). If the average child penalty had stayed constant since 1980, the reduction in Δ_y^o would have been almost twice as large (see dotted pink line). With a constant earnings ratio, instead, overall inequality would have hardly changed (see dash-dotted blue line). Finally, with a fixed share of mothers, the decrease would have been more limited, around 13 percentage points (see dashed red line). In sum, the immense reduction in the overall earnings gap results from the growing earnings potential and the falling fertility rate. By contrast, the growing child penalties prevented a further decrease in overall inequality.

5 The effects of parental leave reforms

From a theoretical perspective, the effect of parental leave policies on gender equality is ambiguous. Proponents argue that such policies can strengthen mothers' labor-market attachment and help them maintain or grow their post-birth earnings. Critics, instead, point out that extended leave can cause long-term income losses (e.g., due to diminished human capital and reduced career progression). Given the patterns in Section 3, suggesting that mothers heavily use parental leave, we now examine whether six German parental leave reforms reduced child penalties and inequality or, conversely, exacerbated them.

5.1 The German parental leave reforms

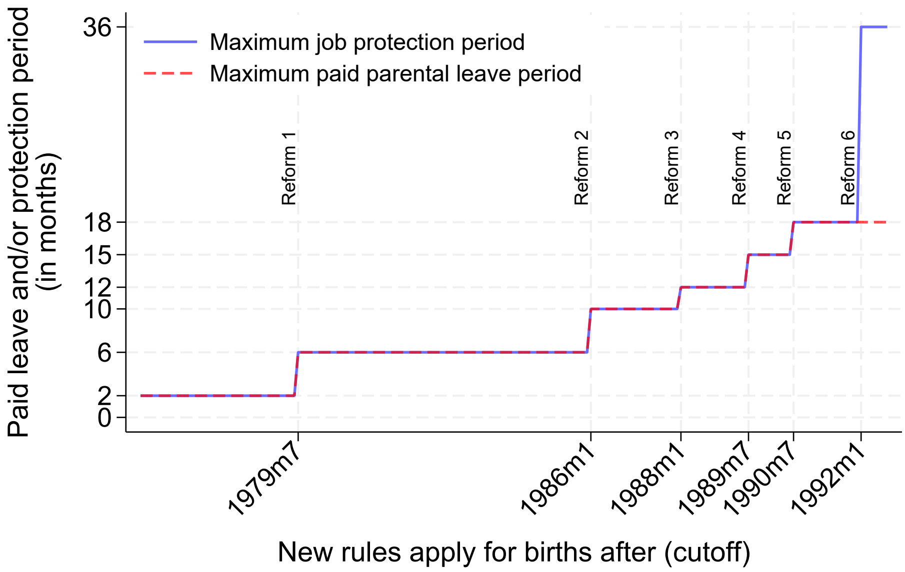
From the 1970s through 1992, the German government rolled out six maternity or parental leave reforms [Schönberg and Ludsteck, 2014]. The six reforms share two central characteristics that enable us to assess their impact on gender inequality. The first is that they all increased the maximum job protection period and (most of them) also the maximum paid leave duration. The second feature is that the reforms introduced a straightforward quasi-experimental variation. The new rules applied to parents who gave birth after specific birthdate cutoffs. We, thus, can compare just affected and just unaffected parents in dynamic regression discontinuity designs to estimate the reform's effects [Schönberg and Ludsteck, 2014; Kleven *et al.*, 2024b]. In the following, we briefly sketch the reforms.

Baseline scenario. Since the mid-1950s, mothers had to take six weeks of mandated maternity leave before and eight weeks after childbirth. During this period, they received wage continuation payments (equivalent to their average net income in the three months prior to birth). Additionally, they benefited from job protection, including the right to return to a similar position and protection against being fired by their employers.

Reforms. Figure 5 summarizes how the reforms changed the maximum paid leave duration (dashed line) and the maximum job protection period (solid line). The first reform in 1979 introduced *voluntary maternity leave*. Specifically, mothers who gave birth on or after July 1, 1979 could take six instead of two months of job-protected paid maternity leave. While the benefits during the mandated leave remained unchanged, benefit payments during voluntary leave amounted to DM 750 (≈ Euro 384).[30] The government implemented a second reform in 1986 that introduced voluntary *parental leave* (cutoff: January 1, 1986). Now, both mothers and fathers became eligible for up to 10 months of job-protected paid leave. However, as pointed out by Schönberg and Ludsteck [2014], only very few fathers took parental leave (less than 1.5% in 1992).

[30] Only women working prior to childbirth became entitled to receive those payments.

Figure 5: Length of work bans, job protection, and maximum paid leave (months)



Notes: This figure summarizes the policy reforms. Specifically, it shows how these reforms changed the job protection periods (solid line) and the maximum paid leave periods (dashed line). Before the first reform in 1979s, German mothers faced a fully paid work ban of eight weeks after childbirth. The first reform in 1979 implemented additional paid parental leave. For six months, mothers who worked pre-birth received DM 750 (≈ Euro 384) monthly. The second reform in 1986 extended the paid leave period to at most 10 months but lowered the monthly payment to DM 600 (≈ Euro 307). Several later reforms jointly expanded paid parental benefits and the job protection period. The last reform only expanded the job protection period.

The reform also lowered the maximum voluntary leave benefits to DM 600 (≈ Euro 307) per month, where they remained until 1993.[31] A series of subsequent reforms further expanded the paid and protected leave period to 12 months (cutoff: January 1, 1988), 15 months (cutoff: July 1, 1989), and 18 months (cutoff: July 1, 1990). The reforms left the benefits payments unchanged. The sixth reform only extended the job protection period to 36 months, while leaving the paid leave duration unaffected at 18 months (cutoff: January 1, 1992). The first five reforms, thus, allow us to identify the effects of joint extensions of paid leave and job protection periods. By contrast, the last reform isolates the pure effect of job protection.

[31] A couple of additional details are noteworthy. First, the new policy entitled all parents to these payments, regardless of their pre-birth employment status. Second, while the policy made the benefit payments from the third to the sixth month after childbirth independent of family earnings, payments from the seventh to the tenth month were means-tested. Specifically, they depended on the annual net family income during the two years preceding childbirth. Notably, almost 85% of women received the full benefit of DM 600 for 10 months.

5.2 Dynamic regression discontinuity designs

We exploit a dynamic regression discontinuity design to estimate the effects of these reforms on mothers' earnings.[32] The design relies on the idea that the birthdate of the first child quasi-randomly places parents under varying parental leave regimes. It, therefore, can act as a treatment assignment variable in a dynamic regression discontinuity (RD) design. Building on this foundation, the intuition of our approach is to estimate "separate" event studies for mothers who delivered their first child just before and just after a specific reform's birthdate cutoff. We then assess if the earnings of affected and unaffected mothers evolve differently across the event time to determine the effects of the reforms.

Estimation approach. We estimate the parameters of an interacted model with OLS that (i) nests those two event studies and (ii) directly identifies differences in the earnings trajectories of mothers who are or are not affected (treated) by one specific reform $z \in \{1979, 1986, \ldots, 1992\}$. The (local) regression uses observations near the cutoff on both sides and assigns greater weight to those closer to it.[33] Using the previous notation, the model reads:

$$Y_{is}^{m} = \sum_{j \neq -12} \alpha_j^{z0} \cdot \mathbb{1}[t=j] + \sum_{j \neq -12} \theta_j^{z} \cdot \mathbb{1}[t=j] \times \mathbb{1}[treat_i = 1] + \mathbf{X_{is}}\eta' + u_{is}, \quad (10)$$

where $\mathbb{1}[treat_i = 1]$ is a dummy variable indicating if mother i had her first child shortly before (untreated) or shortly after (treated) the birthdate cutoff. Moreover, $\mathbf{X_{is}}$ is a vector of control variables (that includes age dummies).

The estimated coefficient $\widehat{\alpha}_t^{z0}$ captures the effect of the first child at event time t for untreated mothers (control group). For treated mothers, the corresponding effect at t is given by $\widehat{\alpha}_t^{z1} = \widehat{\alpha}_t^{z0} + \widehat{\theta}_t^{z}$ (treatment group). Provided that the counterfactual earnings path is smooth when adjusted for controls, the estimates of these coefficients reveal the treatment-status-specific impacts of childbirth on the outcome at event time t. Our main interest, however, lies in understanding how the reform z dynamically impacts mothers' earnings. The estimated coefficient $\widehat{\theta}_t^{z}$ measures this effect at event time t under the assumption that all (unobserved) determinants of the outcome continuously evolve across the birthdate cutoff. As highlighted by Kleven *et al.* [2024b] and Bronson and Sanin [2024], such RDD estimates cleanly capture how the parental leave regime applicable to the first child affects child penalties. They do not, however, identify the

[32] The reforms did not change fathers' leave-taking behavior. Thus, we restrict our analysis to mothers.

[33] Alternatively, we can use polynomial regressions to model the effect of the running variable on the outcome. The key messages are unchanged.

effect of parental leave extensions for subsequent births[34] or the policy's influence on behavior before the first birth [Bronson and Sanin, 2024].[35]

Further details. Several further elements shape our empirical strategy. First, in all figures, we depict our estimated event-time-specific coefficients $\widehat{\alpha}_t^{z0}$, $\widehat{\alpha}_t^{z1}$, and $\widehat{\theta}_t^z$ scaled with the predicted counterfactual earnings $E[\tilde{Y}_{is}^m|t]$ absent children.[36] Hereby, we obtain (i) estimates for child penalties conditional on treatment status and (ii) an estimate of how the reform z affects the size of the *child penalty* at event time t:

$$dP_{t,z} = P_t^{z1} - P_t^{z0} = \frac{\widehat{\theta}_t^z}{E[\tilde{Y}_{is}^m|t]}. \tag{11}$$

Second, to balance bias and variance, we automatically determine the optimal bandwidth around the cutoff using a standard mean square error selector.[37] Our procedure allows for different bandwidths at both sides of the cutoff. Third, we apply two-dimensional weights in our RD design. Our weights combine a triangular kernel (emphasizing observations closer to the threshold) with population weights (adjusting for population representativeness). Fourth, when estimating model (10), we restrict the sample to West German mothers whom we observe at every event time $t \in [-36, 120]$.[38]

[34] Women who give birth to their first child on different sides of the cutoff face identical regimes for any subsequent birth. This feature prevents the identification of effects for subsequent children.

[35] In an ideal RDD setting, the reform at consideration is announced at short notice so that mothers cannot self-select into a particular policy regime. This feature, however, prevents the identification of effects on pre-birth behavior: Mothers on both sides of the cutoff should expect to be subject to the same policy regime, giving them no reason to adjust their behavior before birth.

[36] Specifically, we take the predicted counterfactual earnings that apply in the month prior to the reform. By focusing on the pre-reform month, we (i) establish a consistent baseline for all estimates and (ii) express our estimates relative to a reference group of just-unaffected mothers.

[37] The details of the bandwidth selection procedure are as follows. First, we note that our dynamic model nests static RD models for each event time t. The optimal bandwidth might differ for each of these event times due to variations in data density and volatility around the cutoffs. We, hence, determine these optimal bandwidths separately for each event time. Second, our aim is, however, to implement one single bandwidth across all event times in our interacted model to ensure comparability. We adopt the median of the event-time-specific bandwidths as the standard for all event times. This choice aims at striking a balance between harmonizing the specific needs of each event time with the broader objective of analytical coherence. Third, some of the reforms were implemented in quick succession. We ensure that the chosen bandwidth never includes the next birthdate cutoff.

[38] As before, our model includes event time dummies for each $t \in [-36, 120]$. In addition, to capture broader long-run effects in our decomposition analysis in Subsection 5.4, we integrate a dummy variable encapsulating all periods $t < -36$ and another dummy representing all periods $t > 120$.

5.3 Baseline estimates of reform effects

This subsection discusses the impacts of two specific reforms in detail and, for conciseness, only briefly summarizes the analyses of the other reforms. The Appendix provides more details. We consider this approach to be appropriate as we find very consistent results across all reforms.

Descriptive analysis. Appendix Figure A.21 shows the raw data to exemplify the impacts of the 1979 and 1986 reforms on mothers' earnings descriptively. Each subfigure focuses on a specific event time (i.e., on mothers' earnings at t months before/after birth). Specifically, for each event time, we display mothers' average earnings as a percentage of their earnings 12 months before birth, plotted against the month their first child was born. Four observations stand out. First, before childbirth ($t < 0$), the earnings of mothers evolve smoothly across the birthdate cutoffs (see Appendix Figures A.21a-A.21c). This observation suggests that our identifying assumption holds: Mothers with differential pre-birth earnings do not seem to select one of the two regimes, suggesting that all the other (unobserved) outcome determinants evolve smoothly across the cutoff. Second, mothers' earnings collapse in the two months leading up to childbirth (they take mandatory maternity leave). Third, during the birth month and the subsequent month, earnings remain low without any discontinuity at the birthdate cutoffs (see Appendix Figures A.21d-A.21e). Mothers are still on mandatory leave, rationalizing this result. Fourth, we observe discrete jumps at the policy cutoffs once the reforms have bite (see Appendix Figures A.21f-A.21l). For example, the earnings of mothers who gave birth before July 1979 exhibit a marked increase at event time $t = 2$ (coinciding with the end of the mandatory paid protected leave period). By contrast, the earnings of mothers who deliver in July 1979 or later only recover after $t = 6$ (when the extended voluntary leave period ends). Overall, the figures suggest that the reforms impacted mothers' earnings in the short run (during expanded leave).

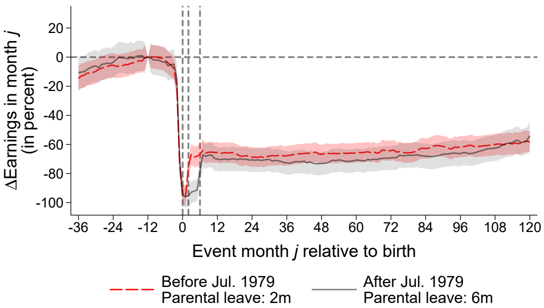
Main results. Figure 6 confirms this descriptive finding. Figure 6a depicts the child penalties for first-time mothers who are treated (solid line) or untreated (dashed line) by the 1979 reform. Figure 6b delineates the corresponding RD reform effects, $d\bar{P}_{t,k}$. The Figures 6c and 6d present analogous estimates for the 1986 reform.

The principal messages are: (i) both reforms caused significant earnings reductions of 12-26 percentage points during the extended leave periods (months 2-6 for the 1979 reform and months 7-10 for the 1986 reform); (ii) yet, we observe no significant long-run effects on earnings. Appendix Figures A.22-A.23 confirm these results for the

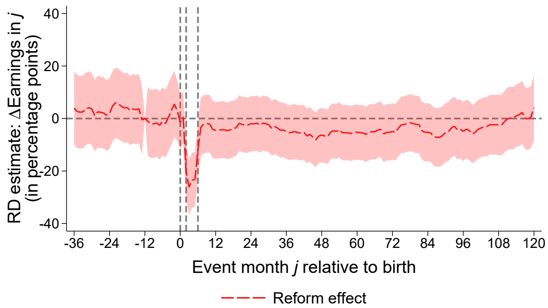
Figure 6: Impacts of two exemplary paid parental leave reforms

Parental leave reform in 1979

(a) Treated versus untreated mothers

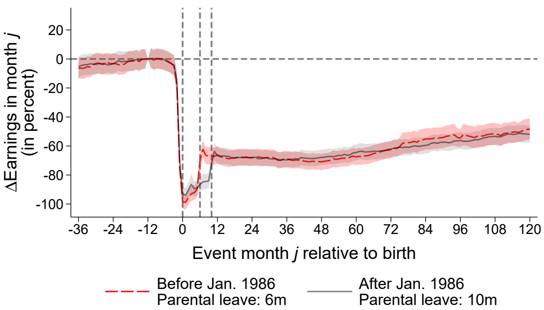


(b) Dynamic RD estimates

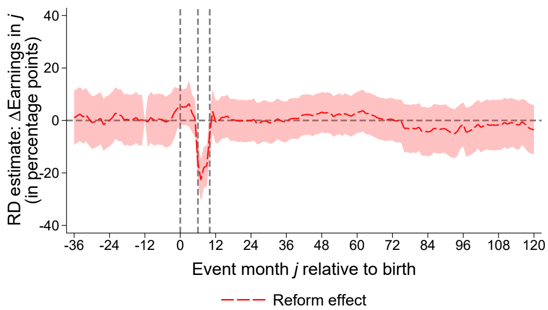


Parental leave reform in 1986

(c) Treated versus untreated mothers



(d) Dynamic RD estimates



Notes: These figures visualize the impacts of paid parental leave reforms in 1979 and 1986 on mothers' earnings trajectories in West Germany. The first row focuses on the 1979 reform and its impact on first-time mothers. Specifically, Figure 6a depicts the percentage impacts of children on the earnings of mothers who are just treated (solid line) or untreated (dashed line) by the reform. By contrast, Figure 6b delineates the corresponding RD reform effects. The second row presents analogous estimates for the 1986 reform.

remaining four parental leave reforms (1988, 1989, 1990, and 1992). As discussed, the six reforms slightly vary in their specific design. However, those differences do not matter much: Neither the bundled reforms nor the reform that only shifted the length of job protection seem to influence mothers' long-term earnings.[39] These results align well with those presented in the papers of Kleven *et al.* [2024b] and Schönberg and Ludsteck [2014], with the latter specifically studying effects of the German reforms on

[39]One explanation for insignificant reform effects might be that our findings represent local effects specific to *compliers*, who might be mothers experiencing minimal career costs from taking leave [Kleven *et al.*, 2024b].

labor market outcomes.[40] The extent to which the German reforms influence overall inequality, however, remains unexplored.

5.4 Parental leave reforms and gender inequality

Thus, our next step is to analyze the impact of the discussed reforms on overall gender inequality. At first glance, one might not expect substantial impacts on overall inequality, given the negligible long-term effects. Yet, given that almost all women get children and leave was extended substantially (from 2 to 36 months), the short-term effects could accumulate and jointly exacerbate gender inequality.

Decomposition methodology. As we have seen in the previous subsection, the German parental leave reforms increased the event-time-specific child penalties, P_t. As a result, the average penalty $\bar{P}_y$ in year y must also have increased, leading to higher child-related and overall inequality in earnings (see equations 4 and 9). Following this logic, our next goal is to quantify these effects of a given reform $z \in \{1979, \ldots, 1992\}$ on child-related and overall inequality in some later year y [see also Kleven *et al.*, 2024b].[41]

We proceed in three steps. The first step is determining how reform z impacts the *average child penalty* $\bar{P}_y$ in year y. In the Appendix, we show that this effect is given by

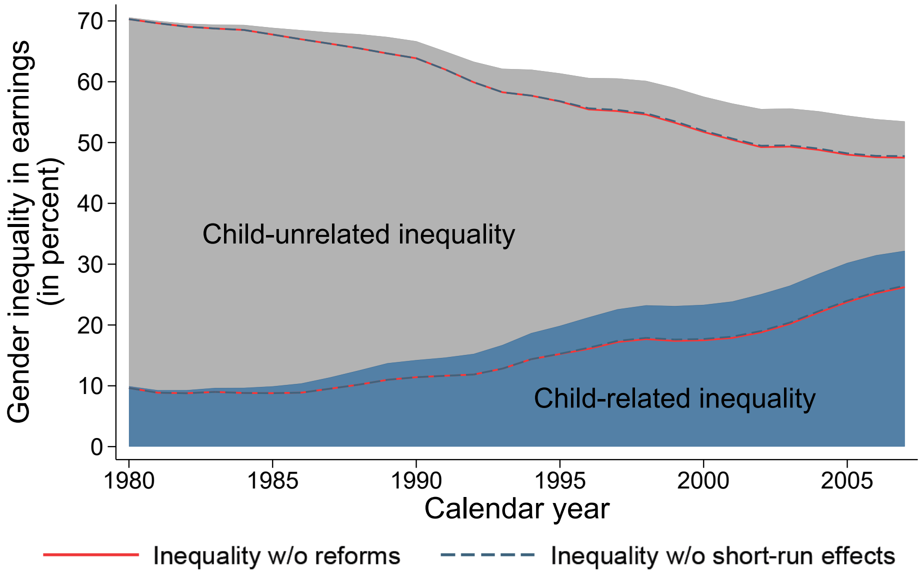
$$d\bar{P}_{y,z} = \frac{E[\widehat{\theta}_t^z \cdot \rho_{t,y}^z \mid y, g_i \geq z]}{E[\tilde{Y}_{is}^m \mid y]}, \tag{12}$$

where $\widehat{\theta}_t^z$ is the reforms' impact in Euros on mothers' earnings at event time t, taken from our RD model (10). The term $\rho_{t,y}^z = E[\tilde{Y}_{is}^m \mid y, t]/E[\tilde{Y}_{is}^m \mid z-1, t]$ is a scaling factor that adjusts for the fact that mothers' counterfactual earnings have grown between the year before the reform and year y. Under the assumption that the percentage effect of the reform on mothers' earnings remains constant over time, the expression in

[40] Kleven *et al.* [2024b] discuss potential reasons why the adverse short-term effects do not translate into longer-term effects. One explanation is that compliance with the reform is higher among mothers who face lower career costs from taking leave. Alternatively, the reforms may (i) only affect earnings before (and not after) birth or (ii) only impact earnings after the birth of subsequent children. As noted earlier, however, RD designs like ours cannot capture these types of effects [Bronson and Sanin, 2024].

[41] In principle, reforms could also affect child-related gender inequality through changing the share of mothers ϕ_y or the counterfactual earnings ratio Ψ_y. However, our RDD cannot investigate the effects on these variables because the underlying decisions (fertility and career planning) are longer-term decisions made before birth [Bronson and Sanin, 2024]. Additionally, Olivetti and Petrongolo [2017] suggest that such reforms only induce small fertility responses. Kleven *et al.* [2024b] conclude that even if there were fertility responses, they would be too minor to shift gender inequality.

Figure 7: Impact of job-protected parental leave reforms on gender inequality



Notes: This figure illustrates how the parental leave reforms between 1979 and 1992 affected overall gender inequality (blue plus gray) and child-related inequality (blue) in earnings in West Germany. We consider two counterfactual scenarios. Scenario 1 (solid red lines) depicts how overall inequality and child-related inequality would have evolved without the leave reforms, accounting for their long-run and short-run effects on child penalties. Scenario 2 (long-dashed blue lines) depicts how the short-run effects changed inequality. The underlying event study models allow for cohort-specific event time coefficients and control for year and age dummies. The sample spans the years from 1980 to 2006 and includes men and women between the ages of 25 and 45.

the numerator of equation (12) represents the expected earnings loss in Euro for the average mother in year y [see Kleven *et al.*, 2024b].[42]

In the second step, we compute the effect of reform z on *child-related inequality* implied by the change in the average child penalty as:

$$d\Delta^r_{y,z} = -\phi_y \cdot d\bar{P}_{y,z} \cdot \Psi_y \times 100. \tag{13}$$

As before, we obtain Ψ_y from equation (8) and take ϕ_y from the data. Consequently, the corresponding level of child-related inequality in the counterfactual without reform z is $\Delta^r_y - d\Delta^r_{y,z}$. In the third step, we can similarly determine *overall inequality* in the same scenario as $\Delta^u_y + \Delta^r_y - d\Delta^r_{y,z}$.

Further details. Three further details of our decomposition strategy are important. First, we introduce two distinct counterfactual scenarios: Scenario 1 assesses how

[42]In the numerator of equation (12), the expectation is taken over the earnings of mothers at all event times t and over all birth cohorts $g_i \geq z$ observed in year y. We condition on $g_i \geq z$ because, of course, the reform can only affect the earnings of mothers with children born after its implementation.

large overall inequality and child-related inequality would have been without the reform. Here, we adjust the realized inequality levels for all long and short-run effects. Scenario 2 only subtracts the short-run impacts during the leave extension phases.[43] Hereby, we derive an estimate for the level of inequality in a world in which the reform produced only short-term effects. By comparing both scenarios, we can analyze which portions of the impacts on inequality stem from short- versus long-run effects. Second, in both scenarios, we assume that the authorities did not implement any of the six reforms. Thus, we analyze how the six reforms jointly impact gender inequality. Third, when constructing our scenarios, we only consider RD reform effects that are statistically significant at the 5% level (as these provide more robust evidence of the reforms' impacts).

Sample: For the reasons discussed in Subsection 4.2, the decomposition sample again includes West German men and women aged 25 to 45. As the government implemented the parental benefit reform in 2007, we mainly focus on the period 1980 to 2006. Appendix Figure A.25 reports similar results for alternative sample definitions.

Main results. Figure 7 demonstrates how the five reforms of paid and job-protected parental leave and the sixth reform of the job protection period jointly affected inequality. To that end, it presents our estimates for the level of overall and child-related inequality in the two scenarios. The solid red lines depict the levels of overall inequality and child-related inequality in a counterfactual world without reforms (Scenario 1). The long-dashed blue lines represent overall and child-related inequality, assuming the reforms triggered only short-term effects (Scenario 2). For completeness, Appendix Figure A.24 additionally shows how the six reforms jointly increased the average child penalty $\bar{P}_y$ in each year y.[44]

Figure 7 conveys two primary insights: The first is that, in line with Figure 6, the short-term effects of the reforms predominantly influence overall gender inequality. The solid red and long-dashed blue lines evolve closely together. Second, even without marked long-term effects, the reforms are a non-negligible driver of overall gender inequality. For example, in 2006, our analysis suggests that overall gender inequality would have been around 6.2 percentage points lower in the absence of the reforms. To put the size of this effect into perspective, note that between 1980 (the year after the first reform) and 2006 (the year before the parental benefit reform), child-related in-

[43] To compute Scenario 2, we set the event-time coefficients for all event times beyond the maximum leave duration (after the reform) to zero.

[44] The reforms, for example, raised the average child penalty in 2006 by about 10 percentage points.

equality increased by almost 22 percentage points (from about 9.9% to approximately 31.5%). Taken at face value, the six reforms account for around 29% of this increase. The rationale why short-run effects can translate into substantial gender inequality is straightforward: The policies extended the time many mothers were absent from the labor force strongly from two to 36 months.

6 The effects of the 2007 parental benefit reform

In 2007, the German government introduced an income-dependent parental benefit. One goal was to counteract long absences from the labor market after childbirth. This section evaluates whether the reform achieved this objective, and how it impacted child penalties (Subsection 6.2) and gender inequality (Subsection 6.3).

6.1 The parental benefit reform

Again, we begin by describing the situation before the reform and then detail the reform's specifics.

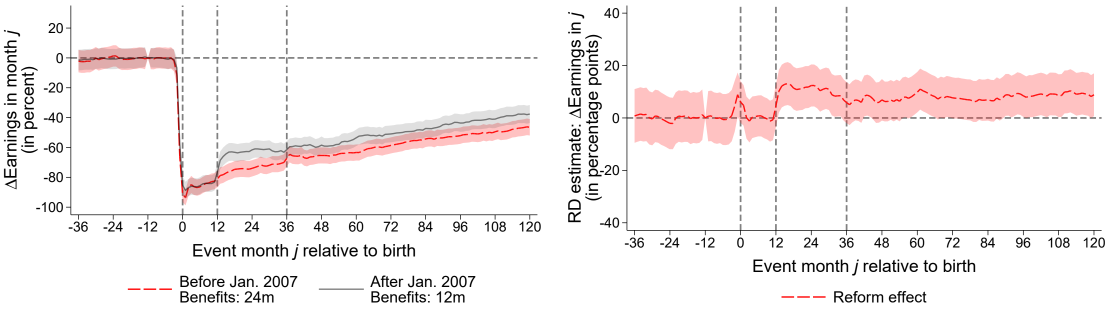
Baseline scenario. Parents of children born before January 1, 2007 received monthly means-tested child-rearing transfers for up to 24 post-birth months. The job protection period stood at 36 months. Families with net (family) incomes below a certain threshold qualified for the full monthly transfer of Euro 300. Above the threshold, the transfer was faced out based on family income. The threshold varied by the number of children, household type (singles/couples), and the time since giving birth. In 2005 and 2006, 77% of parents claimed the maximum transfer for the first six months after childbirth. Subsequent eligibility reviews at 6 and 12 months, along with stricter income limits, reduced the share of parents receiving the full transfer to 47% for months 7 to 12 and 40% for months 12 to 24 [Huebener *et al.*, 2017]. The scheme allowed up to 30 hours of part-time work per week during the benefit period.

Reform. In 2007, the German government replaced the means-tested transfer program with a parental benefit system. Mothers who delivered after January 1, 2007 could receive the new parental benefit payments. This new system shortened the payment period from 24 to 12 months.[45] Under the new scheme, recipients' benefits depended on their average net labor income earned in the 12 months prior to birth.

[45] The benefit scheme granted two additional months to single parents and families where both partners took parental leave for at least two months.

Figure 8: Impacts of the 2007 parental benefit reform

(a) Treated versus untreated mothers **(b)** Dynamic RD estimates



Notes: These figures visualize the impacts of the 2007 parental benefit reform on mothers' earnings trajectories in West Germany. Specifically, Figure 8a depicts the percentage impacts of children on the earnings of mothers who are just treated (solid line) or untreated (dashed line) by the reform. Figure 8b delineates the corresponding RD reform effects.

Recipients without or with very low labor income continued to receive Euro 300 per month. Recipients with higher incomes became eligible for higher benefits, typically amounting to two-thirds of recipients' monthly net earnings before birth, with a cap at Euro 1,800. As a result, most mothers received much higher benefits but for a shorter period after birth. The average payment was 634 Euros per month [Huebener *et al.*, 2017]. By contrast, the reform did not alter the job protection period or the part-time work allowances during the benefit period. The take-up rate for the new parental benefit stands at 96.3% (German Federal Statistical Office).

6.2 Baseline estimates of reform effects

Estimation strategy. Because the child's birthdate effectively assigns parents to different parental benefit regimes in a quasi-random manner, we can again utilize the RD design introduced in Subsection 5.2 to gauge the effects of the 2007 reform on mothers' child penalties. Again, we identify effects for first-time mothers.

Main results. Figure 8 represents the effects of the parental benefit reform. Its structure resembles that of Figure 6: The left figure shows child penalties for treated (solid line) and untreated (dashed line) mothers, and the right examines the reform effects.

We organize the discussion of these effects into four periods. First, the reform left mothers' *pre-birth* earnings trends unchanged (common pre-trends). This observation again suggests that mothers with different characteristics (affecting their earnings paths) did not self-select into different regimes.[46] Second, the reform did not impact child penalties in the *short run* (i.e., during the first twelve months after birth when mothers could receive the new parental benefit). Third, however, the policy change prompted mothers to return to the labor market earlier, and it substantially and significantly reduced *medium-run* child penalties: In each post-birth month 13 (the first month without payments under the new regime) to 36 (the last month with job protection), the post-reform child penalties were 6 to 13 percentage points smaller. There are several explanations for this finding.[47] Fourth, Figure 8 documents *long-run* effects beyond the job protection period. In each event month 37 to 120, child penalties shrank by around 5 to 11 percentage points. The corresponding RD coefficient is either significant or borderline significant at the 5 percent level.

In sum, while we do not find significant long-term effects of the parental leave reforms, the parental benefit reform clearly lowered the child penalties in the long run. This finding aligns with the idea that monetary incentives matter.[48] Indeed, changes in (work) incentives seem to affect mothers' long-run earnings strongly if they, as in our case, occur soon after childbirth [Gruber *et al.*, 2023; Kuka and Shenhav, 2024].

6.3 Parental benefits and gender inequality

In the next step, we investigate the effect of the 2007 reform on overall gender inequality.

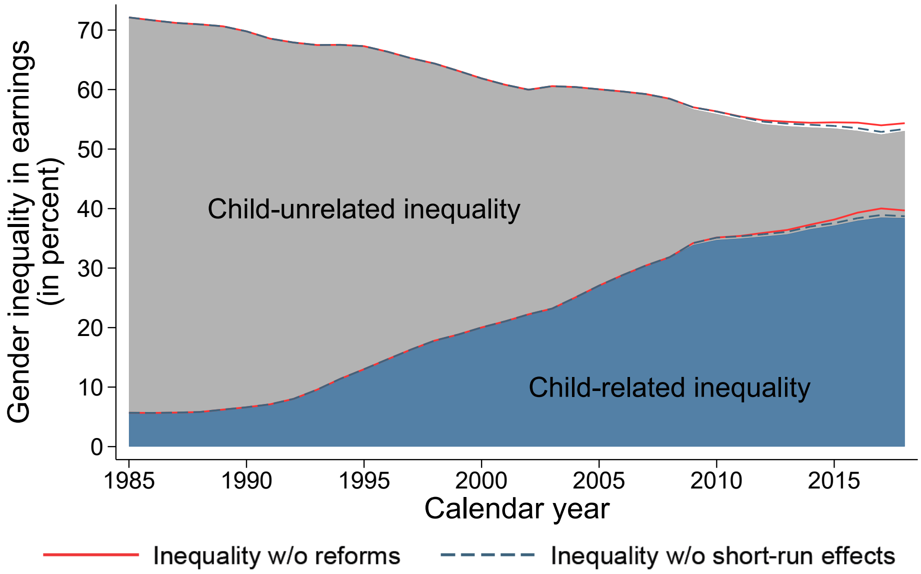
Decomposition methodology. To evaluate the reform's impact on overall and child-related gender inequality, we again use the decomposition framework described in Section 5.4. This analysis contrasts the actual level of gender inequality with two counterfactual scenarios. Scenario 1 estimates inequality in a world without the 2007

[46] Shortly before birth, treated mothers' earnings increase relative to the control group. One explanation is that the benefit amount depends on earnings in the 12 months before birth. Mothers, thus, have an incentive to increase earnings. Glogowsky *et al.* [2024] study this behavior in follow-up work.

[47] One explanation is the incentives from means testing. Before the reform, an increase in earnings in months 12 to 24 could result in a complete loss of parental benefits. This disincentive disappeared after the reform. For some low-income households, income effects may have amplified this channel if the reform made them poorer. A complementary explanation is based on behavioristic preferences: The expiration of benefits after 12 months may have introduced a new reference point for when to return to work.

[48] There are several reasons why monetary incentives could play a significant role. One possibility is that the compliers who react to incentive changes are different ones. For example, changes in monetary incentives might more likely trigger responses by mothers who face greater career setbacks due to extended time out of the labor market (e.g., highly educated women). In our case, higher-earning mothers even face stronger incentives (as the transfers increase in pre-birth income). Another possibility is that such reforms could affect the leave-taking behavior of a larger group of mothers or encourage them to change their leave period more drastically (e.g., from 24 to 12 months).

Figure 9: Impact of the 2007 parental benefit reform on gender inequality



Notes: This figure illustrates how the parental benefit reform in 2007 affected overall gender inequality (blue plus gray) and child-related inequality (blue) in earnings in West Germany. We consider two counterfactual scenarios. Scenario 1 (solid red lines) depicts how overall inequality and child-related inequality would have evolved without the leave reforms, accounting for their long-run and short-run effects on child penalties. Scenario 2 (long-dashed blue lines) depicts how the short-run effects changed inequality. The underlying event study models allow for cohort-specific event time coefficients and control for year and age dummies. The sample spans the years from 1985 to 2018 and includes men and women between the ages of 30 and 50.

parental benefit reform. Here, we adjust mothers' realized earnings by all significant short-run, medium-run, and long-run effects. Scenario 2, instead, considers only short-run and medium-run effects (up to $t = 36$).

Sample. Compared to Section 5, the following analysis relies on a different sample. The reason is that the effects of the parental benefit reform unfold over the long run (ten years). If we maintained our standard sample restriction (1980 to 2013), we would have missed these long-run effects in our decomposition analysis (as the reform was implemented in 2007). Thus, we expand our decomposition to later years. Our new sample covers the years 1985 to 2018. To ensure that we capture the full age range across all considered years, we also need to adjust the age restriction: We now include all individuals aged 30 to 50. Appendix Figure A.27 presents the corresponding results for our standard sample.

Main results. Figure 9 illustrates that the reform prevented a further increase in child-related and overall inequality. After 2006, overall gender inequality remained almost constant, while child-related inequality increased by 9.6 percentage points (from 28.8% to 38.4%). Without the parental benefit reform, instead, child-related inequality would have further increased by an additional 1.3 percentage points according to our estimates (see Scenario 1 depicted by the solid red lines). Put differently, without the reform, this part of inequality would have grown by 13.6% more than it did with the reform. This pattern translates into an equivalent increase in overall inequality. The comparison between Scenario 1 (solid red lines) and Scenario 2 (long-dashed blue lines) further demonstrates the crucial role of long-run effects: Both medium- and long-run effects contributed, but the long-run effects had a more substantial impact on reducing inequality. For completeness, Appendix Figure A.26 additionally shows that the reform reduced the average child penalty $\bar{P}_y$ in 2018 by around 2 percentage points. We conclude that the parental benefit reform countered the rise in child-related inequality.

Discussion. Notably, the actual long-run impacts of the reform on inequality are likely even bigger (see Appendix C.3 for a formal discussion). Our analysis highlights two main reasons why the policy impact will increase over time. The first reason is rather mechanical: Because only mothers who gave birth in 2007 or later are treated, the share of treated mothers increases over time until all mothers have given birth under the new regime. Thus, we observe a larger share of mothers who changed their behavior due to the reform in later years. Moreover, as time passes, the average age of treated mothers increases. In 2007, for example, the reform only affected mothers with children below one year. By contrast, in 2022, all mothers with children below 16 gave birth under the new regime. This observation leads us to the second reason why the policy impact increases over time: Mothers with older children have higher potential earnings due to experience effects (captured by our age dummies). If we, for example, assume that the eight percentage point drop in the child penalty stays constant after event time $t = 120$, the reform's impact on earnings grows in absolute units as the child ages. Hence, the impact on $\bar{P}_y$ is larger for mothers with older children. We conclude in Appendix C.3 that likely less than half of the long-run reform effect on inequality was realized by 2018.

7 Conclusion

This paper demonstrates how (i) parenthood and (ii) family policies have shaped gender inequality in Germany over the past fifty years. We contribute to the literature by presenting three key sets of results.

The first set of results documents the evolution of child penalties in Germany. As a benchmark, we start with reporting child penalties for mothers who delivered their first child after the German reunification. This analysis verifies the substantial and persistent impact of parenthood on mothers' earnings in recent times. Next, we demonstrate how child penalties have developed since the 1960s. The key insight of our work is that child penalties grew over time: German mothers who delivered their first child in the 2000s experienced much larger percentage losses in earnings than mothers giving birth in the 1960s. This pattern is general in that we observe it across various subgroups (e.g., within all educational levels). In light of these results, a key question is how much of the overall earnings inequality is explained by parenthood.

The second set of results provides insights into this question through decomposition analyses. We first apply the standard approach of Kleven *et al.* [2019b] to our setting and decompose gender inequality in earnings into a child-related and a child-unrelated component. For comparison, note that in Germany, overall gender inequality declined: For example, women aged 25 to 45 earned about 71% less than men in 1980 and about 54% less in 2013. At the same time, the share of gender inequality that we can relate to children increased by a factor of close to five (from about 14% to almost 64%). We then introduce an extended decomposition approach, revealing that, in addition to rising child penalties, a second fundamental force contributed to the increase in child-related inequality: a rise in potential earnings. Mothers who gave birth in more recent decades had higher potential earnings, leading to greater losses when they exited the labor market. This result is crucial because, from a normative perspective, only the increase in child penalties appears worrying, while the rise in potential earnings is arguably welcomed.

In the third set of results, we show that German family policies shaped gender inequality to a non-negligible extent. For example, six expansions of job-protected parental leave between 1979 and 1992 accounted for 29% of the rise in child-related inequality between 1980 and 2006. Interestingly, short-run effects that manifest during the expanded leave period dominate this impact on overall inequality. The core logic for how short-run effects accumulate into substantial gender inequality is that the policies impacted most women, as most have children, and removed them from the workforce for a considerable time (36 months instead of two). By contrast, we find that the 2007 parental benefit reform marked a successful intervention by the German government to counteract the trend of increasing child-related inequality. The reform motivated mothers to choose higher earnings both in the medium run and in the long run. According to our estimates, without the reform, child-related inequality would have increased by 13.6% more than it did with the reform.

These findings carry important implications for policy-making. They show that sidelining mothers after childbirth becomes increasingly costly as their earnings potential grows. They also highlight how family policies can influence gender inequality, with extended parental leave increasing and short-run benefits helping to reduce it. While our paper, thus, offers many new insights, numerous questions remain. A key area for future research is understanding all the ways through which policy affects behavior. Do policies shift gender norms, leading to lasting changes in gender roles? Do they create anticipation effects or even alter mothers' potential earnings trajectories? These are pressing issues for policymakers, and much work remains to be done.

References

AKERLOF, G. A. and KRANTON, R. E. (2000). Economics and Identity. *Quarterly Journal of Economics*, **115** (3), 715–753.

ANDRESEN, M. E. and NIX, E. (2022). What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples. *Journal of Labor Economics*, **40** (4), 971–1004.

— and — (2023). *You Can't Force Me Into Caregiving: Paternity Leave and the Child Penalty*. Discussion Paper.

ANGELOV, N., JOHANSSON, P. and LINDAHL, E. (2016). Parenthood and the Gender Gap in Pay. *Journal of Labor Economics*, **34** (3), 545–579.

ANGRIST, J. D. and EVANS, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*, **88** (3), 450–477.

BAILEY, M., BYKER, T., PATEL, E. and RAMNATH, S. (2024). *The Long-Run Effects of Californiaâs Paid Family Leave Act on Womenâs Careers and Childbearing: New Evidence from a Regression Discontinuity Design and US Tax Data*. Working paper.

BAUERNSCHUSTER, S. and SCHLOTTER, M. (2015). Public Child Care and Mothers' Labor Supply–Evidence from Two Quasi-Experiments. *Journal of Public Economics*, **123**, 1–16.

BECKER, S. O., MERGELE, L. and WOESMANN, L. (2020). The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism. *Journal of Economic Perspectives*, **34** (2), 143–171.

BENSNES, S., HUITFELDT, I. and LEUVEN, E. (2023). *Reconciling Estimates of the Long-Term Earnings Effect of Fertility*. Discussion Papers 1004, Statistics Norway.

BERTRAND, M., GOLDIN, C. and KATZ, L. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, **2** (3), 228–255.

BLAU, F. D. and KAHN, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, **55** (3).

BOELMANN, B., RAUTE, A. and SCHÖNBERG, U. (2024). Wind of Change? Cultural Determinants of Maternal Labor Supply. *American Economic Journal: Applied Economics*, **forthcoming**.

BÖNKE, T., CORNEO, G. and LÜTHEN, H. (2015). Lifetime Earnings Inequality in Germany. *Journal of Labor Economics*, **33** (1), 171–208.

BRONSON, M. A. and SANIN, D. (2024). Female Labor Supply, Fertility and Parental Leave Policy Design, mimeo.

BUSSE, A. and GATHMANN, C. (2020). Free Daycare Policies, Family Choices and Child Development. *Journal of Economic Behavior & Organization*, **179**, 240–260.

CAMPA, P. and SERAFINELLI, M. (2019). Politico-Economic Regimes and Attitudes: Female Workers Under State Socialism. *Review of Economics and Statistics*, **101** (2), 233–248.

CANAAN, S., LASSEN, A., ROSENBAUM, P. and STEINGRIMSDOTTIR, H. (2022). *Maternity Leave and Paternity Leave: Evidence on the Economic Impact of Legislative Changes in High-Income Countries*. IZA Discussion Paper 15129, IZA.

CORTÉS, P. and PAN, J. (2023). Children and the Remaining Gender Gaps in the Labor Market. *Journal of Economic Literature*, **61** (4), 1359–1409.

GATHMANN, C. and SASS, B. (2018). Taxing Childcare: Effects on Childcare Choices, Family Labor Supply, and Children. *Journal of Labor Economics*, **36** (3), 665–709.

GLOGOWSKY, U., GROSENICK, A., HANSEN, E., KOCH, L., PEICHL, A. and SACHS, D. (2024). The Impact of Parental Leave Benefits on Pre-Birth Earnings, mimeo.

GRUBER, J., KOSONEN, T. and HUTTUNEN, K. (2023). *Paying Moms to Stay Home: Short and Long Run Effects on Parents and Children*. NBER Working Paper 30931, National Bureau of Economic Research, Cambridge, MA.

HUBER, K. and ROLVERING, G. (2023). *Public Child Care and Mothers' Career Trajectories*. IZA Discussion Paper 16433, IZA.

HUEBENER, M., KUEHNLE, D. and SPIESS, K. (2017). *Paid Parental Leave and Child Development: Evidence from the 2007 German Parental Benefit Reform and Administrative Data*. DIW Berlin Discussion Paper 1651.

JESSEN, J. (2022). Culture, Children and Couple Gender Inequality. *European Economic Review*, **150**, 104310.

—, SCHWEIGHOFER-KODRITSCH, S., WEINHARDT, F. and BERKES, J. (2024). *Separate Housework Spheres*. IZA Discussion Paper 17134, IZA.

KLEVEN, H. (2023). *The Geography of Child Penalties and Gender Norms: Evidence from the United States*. NBER Working Paper 30176, National Bureau of Economic Research, Cambridge, MA.

—, LANDAIS, C. and LEITE-MARIANTE, G. (2024a). The Child Penalty Atlas. *The Review of Economic Studies*, **forthcoming**.

—, —, POSCH, J., STEINHAUER, A. and ZWEIMÜLLER, J. (2019a). Child Penalties Across Countries: Evidence and Explanations. *American Economic Review (Papers and Proceedings)*, **109**, 122–126.

—, —, —, — and ZWEIMÜLLER, J. (2024b). Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation. *American Economic Journal: Economic Policy*, **forthcoming**.

—, — and SØGAARD, J. E. (2019b). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, **11** (4), 181–209.

KRAPF, M., ROTH, A. and SLOTWINSKI, M. (2020). *The effect of childcare on parental earnings trajectories*. CESifo Working Paper 8764, CESifo.

KREYENFELD, M. and GEISLER, E. (2006). Müttererwerbstätigkeit in Ost-und Westdeutschland. *Zeitschrift für Familienforschung*, **18** (3), 333–360.

KUKA, E. and SHENHAV, N. (2024). Long-Run Effects of Incentivizing Work After Childbirth. *American Economic Review*, **114** (6), 1692–1722.

KUZIEMKO, I., PAN, J., SHEN, J. and WASHINGTON, E. (2018). *The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?* NBER Working Paper 24740, National Bureau of Economic Research.

LUNDBORG, P., PLUG, E. and RASMUSSEN, A. W. (2017). Can Women Have Children and a Career? IV Evidence from IVF Treatments. *American Economic Review*, **107** (6), 1611–1637.

—, — and — (2024). *Is There Really a Child Penalty in the Long Run? New Evidence from IVF Treatments*. IZA Discussion Paper 1004, IZA.

MELENTYEVA, V. and RIEDEL, L. (2023). *Child Penalty Estimation and Mothers' Age at First Birth*. ECONtribute Discussion Paper 266, ECONtribute.

MISCHLER, F. (2021). *Verdienstunterschiede zwischen Männern und Frauen*. WISTA 4, Statistisches Bundesamt (Destatis).

OLIVETTI, C. and PETRONGOLO, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, **8** (1), 405–434.

— and — (2017). The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries. *Journal of Economic Perspectives*, **31** (1), 205–230.

SCHÖNBERG, U. and LUDSTECK, J. (2014). Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth. *Journal of Labor Economics*, **32** (3), 469–505.

WALDFOGEL, J. (1998). Understanding the 'Family Gap' in Pay for Women with Children. *Journal of Economic Perspectives*, **12** (1).

Web Appendix

This Web Appendix provides additional material discussed in the unpublished manuscript “Parenthood, Family Policies, and Gender Inequality in Germany: From the 1960s to Today” by Ulrich Glogowsky, Emanuel Hansen, Dominik Sachs, Timm Bönke, and Holger Lüthen.

A Empirical Appendix: Matching strategy

Technically, the optimal comparison group would measure how mothers' earnings evolve around childbirth when they never had children. Our objective is to closely approximate such a counterfactual using childless women.

Matching. A key challenge in this endeavor is matching childless women to mothers in a way that ensures the comparability of both groups. Kleven *et al.* [2019b] proposes a matching strategy that (i) uses *all* childless women as a comparison group and (ii) mainly matches childless women to mothers based on age. One intuitive way to think about the strategy is that it assigns hypothetical birth events to all childless women. The assignment happens in a way that ensures that the age distribution of childless women at this hypothetically assigned birth event closely resembles the actual age-at-first distribution for mothers.

The details of the approach are as follows: We draw a hypothetical age for each childless woman at which she would have had her first child (if she had one) from the observed distribution of actual mothers' age at first birth. To account for heterogeneity across space and time, we conduct separate draws within cells of mothers' regions of residence (East versus West) and women's birth cohorts. More specifically, we model the true age-at-first-birth distribution of cohort-k mothers living in region r with a log-normal distribution

$$\textit{age at first birth}_{kr} \sim \mathscr{L}\mathscr{N}(\mu_{kr}, \sigma^2_{kr}),$$

where we obtain estimates for the mean μ_{kr} and variance σ^2_{kr} from the observed ages at first birth within each kr cell. We then draw a hypothetical age at first birth for each childless woman from this distribution and calculate the corresponding pseudo event date.[1]

Estimating event studies around pseudo events. We can then estimate model (1) for the sample of childless women c around the simulated birth events. The estimated coefficients for the hypothetical event dummies $\widehat{\alpha}^c_t$ and those of $\widehat{P}^c_t$ reflect general patterns in women's earnings (i) that evolve around the time mothers typically get children and (ii) that are not filtered out by age and year dummies. Alternatively, we

[1] As robustness checks, we can match mothers and childless women additionally based on their educational level. In practice, we then conduct separate draws within cells of mothers' regions of residence (East versus West), their birth cohorts, and their educational level. The results remain unchanged.

can also use these matched women as a comparison group in a difference-in-difference framework.

Child-related earnings gaps. Building on the previous event-study estimates, we can estimate an alternative child penalty — relative to childless women — as:

$$\widehat{P}_t^{relativ,childless} = \frac{\widehat{\alpha}_t^c - \widehat{\alpha}_t^m}{E[\tilde{Y}_{ist}^m|t]}.$$

Intuitively, $\widehat{P}_t^{relativ,childless}$ reflects by how many percent mothers' earnings at event time t fall behind those of childless women because of children. Kleven *et al.* [2019b] highlight the underlying identifying assumption allowing a causal interpretation of this parameter. We must assume parallel counterfactual earnings trends for mothers and childless women over the event time, conditional on year and age dummies. We have tested the robustness of our results against using this alternative empirical strategy. We find that $\widehat{P}_t^{relativ,childless}$ and $\widehat{P}_t$ are almost indistinguishable. The reason is that, as shown in our main figures, the estimates of $\widehat{\alpha}_t^c$ are very close to zero.

B Data Appendix: The German Taxpayer Panel

Dataset. We utilize the German Taxpayer Panel as our secondary data source. The German tax authorities collect this administrative tax dataset, which the German Federal Statistical Office then provides to researchers. The primary purpose of these data is to calculate the income tax, so the dataset includes all necessary information for this calculation. We gained access to a random 5% sample of the full dataset, covering the period from 2001 to 2014. The unit of observation is "a taxpayer" (i.e., either an individual or a couple filing jointly). Several features render this dataset ideal for our robustness check. First, it provides detailed income information. Second, it includes basic demographic data, such as children's birthdates and the number of children. This enables us to analyze earnings trajectories around the time of childbirth. Third, it covers not only employees but also civil servants and self-employed individuals. Fourth, most importantly, this dataset allows us to match children to their fathers. The dataset, thus, allows us to assess how children impact fathers' income, an analysis impossible in the pension data. Despite these useful features, we do not use this dataset for our main analysis because it (i) covers a much shorter period and (ii) lacks monthly information.

Estimation Sample. We focus on individuals who gave birth between 2003 and 2010, and we track them over two years before and four years after birth. Similar to our main analysis, we estimate child penalties using a balanced sample. Consequently, we only include individuals in our regression whom we observe for two pre-birth and four post-birth periods. Appendix Figure A.6 shows the results.

C Formal Appendix

C.1 Decomposition of gender inequality in earnings

In this section, we provide further details on how we decompose overall gender inequality and child-related inequality in earnings.

Consider a population of men and women (with and without children) for whom we observe labor earnings over a sequence of months in a set of years $\mathscr{Y} = \{y_1, \ldots, y_T\}$. In year $y \in \mathscr{Y}$, we define the raw gender gap in earnings – henceforth referred to as *overall gender inequality in earnings* – as

$$\Delta_y^o \quad = \quad \frac{E[Y_{is}^{men}|y] - E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100,$$

where Y_{is}^{men} (Y_{is}^{women}) is the observed labor income of man (woman) i in calendar month s. To compute Δ_y^o, we take the difference between the average earnings of men and those of women in all months belonging to year y. We divide this difference by the average earnings of men, thereby expressing inequality as the percentage number by which women's earnings fall behind those of men.

Next, let $\tilde{Y}_{is}^{men}$ ($\tilde{Y}_{is}^{women}$) denote the potential earnings of man (woman) i in a state of the world where he (she) has no children. We can then decompose overall inequality as follows:

$$\begin{aligned} \Delta_y^o \quad = \quad & \frac{E[\tilde{Y}_{is}^{men}|y] - E[\tilde{Y}_{is}^{women}|y]}{E[Y_{is}^{men}|y]} \times 100 \\ & + \frac{\left(E[\tilde{Y}_{is}^{women}|y] - E[Y_{is}^{women}|y]\right) - \left(E[\tilde{Y}_{is}^{men}|y] - E[Y_{is}^{men}|y]\right)}{E[Y_{is}^{men}|y]} \times 100, \end{aligned}$$

where, on the right-hand side, the term in the first line is *child-unrelated inequality* Δ_y^u – the part of inequality that would arise without any children – and the term in the second line is *child-related inequality* Δ_y^r – the part of inequality that can be attributed to parenthood.

Under the assumption that men's earnings are not affected by children – neither for fathers nor for childless men – the terms $E[\tilde{Y}_{is}^{men}|y]$ and $E[Y_{is}^{men}|y]$ equal each other. Then, both components of inequality simplify to the expressions in equation (4) in the main text,

$$\begin{aligned}
\Delta_y^u &= \frac{E[Y_{is}^{men}|y]-E[\tilde{Y}_{is}^{women}|y]}{E[Y_{is}^{men}|y]}\times 100,\\
\Delta_y^r &= \frac{E[\tilde{Y}_{is}^{women}|y]-E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]}\times 100.
\end{aligned}$$

In the next step, we note that, for childless women, observed and potential earnings necessarily coincide as well. For mothers, by contrast, $\tilde{Y}_{is}^{women}$ are *counterfactual* earnings that are taken to differ from the observed earnings Y_{is}^{women}. For estimates of these counterfactual earnings, we exploit the event study approach specified in equation (1). To allow for changes in the data generation process over time, however, we segment our main sample according to the year of birth of a mother's first child (e.g., the child cohorts) and run a separate event study for each segment. Hence, we take the earnings of a mother i with a first child born in year g_i as determined by her event time t, her age k, the calendar year y and an idiosyncratic error term:

$$\begin{aligned}
Y_{is}^{m} &= \widehat{\alpha}_t^{g_i}+\widehat{\beta}_k^{g_i}+\gamma_y^{g_i}+\widehat{u}_{is}\\
&= \widehat{\alpha}_t^{g_i}+\tilde{Y}_{is}^{m}+\widehat{u}_{is},
\end{aligned}$$

where $\widehat{\alpha}_t^{g_i}$ is the estimated effect of motherhood on her earnings, and the sum of the remaining terms represents the best-available estimate of her counterfactual earnings. A subtle aspect is that, a priori, we would predict her potential to equal $\tilde{Y}_{is}^{m}=\widehat{\beta}_k^{g_i}+\gamma_y^{g_i}$. Upon observing the realization Y_{is}^{m}, however, we possess an estimate of the idiosyncratic component $\widehat{u}_{is}$. By assumption, the latter is not driven by her motherhood; we hence assign it to the potential (child-unrelated) part of her earnings.[2]

Denoting by ϕ_y the share of mothers among all women in year y, child-related inequality can be written as

$$\begin{aligned}
\Delta_y^r &= \frac{E[\tilde{Y}_{is}^{women}|y]-E[Y_{is}^{women}|y]}{E[Y_{is}^{men}|y]}\times 100\\
&= \phi_y\cdot\frac{E[\tilde{Y}_{is}^{m}+\widehat{u}_{is}|y]-E[Y_{is}^{m}|y]}{E[Y_{is}^{men}|y]}\times 100\\
&= -\phi_y\cdot\frac{E[\widehat{\alpha}_t^{g_i}|y]}{E[Y_{is}^{men}|y]}\times 100,
\end{aligned}$$

see equation (6) in the main text. In the last line, the expectation in the numerator is

[2] Note that $E[\widehat{u}_{is}\mid g_i]$ equals zero by construction, while $E[\widehat{u}_{is}\mid y]$ does not necessarily.

taken over event times and child cohorts of all mothers observed in year y. Formally,

$$E[\widehat{\alpha}_t^{g_i}|y] = \sum_{g_i}\sum_t \frac{\phi_{t,y}^{g_i}}{\phi_y} \cdot \widehat{\alpha}_t^{g_i},$$

where $\phi_{t,y}^{g_i}$ is the share of mothers from child cohort g_i who are at event time t among all women observed in year y. Note that $\phi_y = \sum_{g_i}\sum_t \phi_{t,y}^{g_i}$. Also, note that we include in the share of mothers in y those women who will have a child in the following years, i.e., who are at pre-birth event times $t < 0$.[3]

In the final step, to derive the expressions in (9), we expand the fraction in equation (6) with the expected counterfactual earnings of mothers, $E[\tilde{Y}_{is}^m \mid y]$,

$$\begin{aligned}
\Delta_y^r &= -\phi_y \cdot \frac{E[\widehat{\alpha}_t^{g_i}|y]}{E[Y_{is}^{men}|y]} \cdot \frac{E[\tilde{Y}_{is}^m \mid y]}{E[\tilde{Y}_{is}^m \mid y]} \times 100 \\
&= -\phi_y \cdot \frac{E[\widehat{\alpha}_t^{g_i}|y]}{E[\tilde{Y}_{is}^m \mid y]} \cdot \frac{E[\tilde{Y}_{is}^m \mid y]}{E[Y_{is}^{men}|y]} \times 100 \\
&= -\phi_y \cdot \bar{P}_y \cdot \Psi_y \times 100.
\end{aligned}$$

The expression in the last line results from using the average child penalty $\bar{P}_y$ and the ratio Ψ_y of potential earnings of mothers relative to men, as defined in (7) and (8).

C.2 Reform effects on gender inequality in earnings

This section explain how we compute the effects of the German policy reforms on child-related and overall inequality and the predicted evolution ob both measures in a counterfactual state of the world without these reforms. Our approach follows Kleven *et al.* [2024b], adapted for the extended decomposition of child-related inequality Δ_y^r.

We index the reforms by the year of their implementation, $z \in \mathcal{Z} = \{1979, 1986, 1988, 1989, 1990, 1992\}$. According to equation (9), *child-related inequality* can be written as

$$\Delta_y^r = -\phi_y \cdot \bar{P}_y \cdot \Psi_y \times 100,$$

where the average child penalty is given by

$$\bar{P}_y = \frac{E[\widehat{\alpha}_t^{g_i} \mid y]}{E[\tilde{Y}_{is}^m \mid y]}.$$

[3] But, as the child effects pre-birth are insignificant and hardly distinguishable from zero in all but the very last months before birth, this convention does not turn out to have an effect on the level of child-related inequality.

In Subsection 5.3, we show that each reform z significantly affected the absolute earnings loss $\widehat{\alpha}_t$ at certain event times t. More precisely, the estimated effect of reform z on the earnings loss at event time t is $\widehat{\theta}_t^z$. Our goal is to compute the implied effects on the average child penalty $\bar{P}_y$ and how they translate into child-related inequality Δ_y^r and overall inequality Δ_y^o.

In the first step, we quantify the effects that a reform implemented in year z would have had on the average child penalty in some later year y. Under the assumption that the percentage effect of each reform on mothers' earnings stay constant over time, see Kleven *et al.* [2024b], this effect follows as

$$
\begin{aligned}
d\bar{P}_{y,z} &= \frac{1}{E[\tilde{Y}_{is}^m \mid y]} \cdot \sum_t \sum_{g_i \geq z} \frac{\phi_{t,y}^{g_i}}{\phi_y} d\alpha_t^z \rho_{t,y}^z \\
&= \frac{1}{E[\tilde{Y}_{is}^m \mid y]} \cdot \sum_t \sum_{g_i \geq z} \frac{\phi_{t,y}^{g_i}}{\phi_y} \widehat{\theta}_t^z \rho_{t,y}^z \\
&= \frac{E[\widehat{\theta}_t^z \cdot \rho_{t,y}^z \mid y, g_i \geq z]}{E[\tilde{Y}_{is}^m \mid y]}. \qquad (14)
\end{aligned}
$$

To understand where this expression comes from, note first that the effects of each reform depend on the event time t. Hence, we need to account for how many mothers are at event time t in year y. Second, each reform only affects mothers of children born after a certain birthdate cutoff. Hence, the effects are restricted to mothers with child cohorts $g_i \geq z$. Third, the potential earnings of mothers grow from $E[\tilde{Y}_{is}^m \mid z-1, t]$ just before the reform was implemented to $E[\tilde{Y}_{is}^m \mid y, t]$ in calendar year y. The scaling factor $\rho_{t,y}^z$ ensures that the percentage effect of the reform on mothers' potential earnings stays constant over time,

$$
\begin{aligned}
\frac{\theta_t^z}{E[\tilde{Y}_{is}^m \mid z-1, t]} &= \frac{\theta_t^z \rho_{t,y}^z}{E[\tilde{Y}_{is}^m \mid y, t]} \\
\Leftrightarrow \rho_{t,y}^z &= \frac{E[\tilde{Y}_{is}^m \mid y, t]}{E[\tilde{Y}_{is}^m \mid z-1, t]}.
\end{aligned}
$$

Finally, note that, in (14), we set to zero all coefficients θ_t^z that are not statistically significant at the 5% level in our RD analysis.

In the second step, we compute the effect of reform z on *child-related inequality* implied by the change in the average child penalty as:

$$
d\Delta_{y,z}^r = -\phi_y \cdot d\bar{P}_{y,z} \cdot \Psi_y \times 100.
$$

Consequently, the corresponding level of child-related inequality in the counterfactual without reform z is $\Delta_y^r - d\Delta_{y,z}^r$. Likewise, we can predict the counterfactual inequality in year y in a counterfactual scenario without multiple reforms by subsequently subtracting the predicted effects of each reform from the observed level of inequality. For example, in 1991, the predicted level of child-related gender inequality in earnings without any of the parental leave reforms implemented up to this year is given by

$$\Delta_{1991}^r - \sum_{z \in \mathscr{Z} \setminus \{1992\}} d\Delta_{1991,z}^r.$$

For the third step, recall that overall inequality in earnings is given by $\Delta_y^o = \Delta_y^r + \Delta_y^u$. Assuming that child-unrelated inequality is not affected by the parental leave reforms, we can determine *overall inequality* in the counterfactual scenario without reform z as $\Delta_y^u + \Delta_y^r - d\Delta_{y,z}^r$.

C.3 Long-run effects of parental leave benefit reform

This section discusses formally which share of the long-run effects that the parental-leave benefit reform of 2007 has on child-related inequality have been realized in our last sample year 2018.

Recalling (12) and using $\rho_{t,y}^z = E[\tilde{Y}_{is}^m \mid y, t]/E[\tilde{Y}_{is}^m \mid z-1, t]$, one can express the impact of this reform on the average child penalty in year y as

$$d\bar{P}_y = \frac{E[\widehat{\theta}_t^{2007} \cdot \rho_{t,y}^{2007} \mid y, g_i \geq 2007]}{E[\tilde{Y}_{is}^m \mid y]} = \frac{E[dP_t \tilde{Y}_{is}^m \mid y, g_i \geq 2007]}{E[\tilde{Y}_{is}^m \mid y]}.$$

For notational simplicity, we consider the case of annual data as opposed to monthly data in the following. Hence, the above equation can be written as

$$d\bar{P}_y = \frac{E[dP_t \tilde{Y}_{iy}^m \mid g_i \geq 2007]}{E[\tilde{Y}_{iy}^m]}.$$

To work out transparently how $d\bar{P}_y$ changes over time y, we make several assumptions.

Assumption 1: *Assume that all mothers have their first child at age* 30.

This assumption makes the following derivation simple because 30 is also the lower bound of the age span that we consider ($30-50$). More important, 30 was approximately the average age of German mothers at first birth for the decade $2010-2020$.

Under this assumption, we can write

$$d\bar{P}_y = \frac{\sum_{a=30}^{50} f(a|y) dP_{a-30} \mathbb{1}_{ya} \bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|y)\bar{Y}_{y,a}}.$$

where $f(a|y)$ is the share of mothers in year y that are of age a, hence $\sum_{a=30}^{50} f(a|y) = 1$. Event time (recall that here it is expressed in years as opposed to months for simplicity) is given by $a-30$. $\bar{Y}_{y,a}$ denotes counterfactual earnings of mothers of age a in calendar year y. Finally, $\mathbb{1}_{ya}$ takes the value one if mothers of age a have been treated already in year y.

Recall that the reform was implemented in 2007, i.e. $y = 07$. Then in this calendar year 2007, only the 30 year old mothers are affected as for them the reform just applies. Hence, we have

$$d\bar{P}_{07} = \frac{f(30|07) dP_0 \bar{Y}_{07,30}}{\sum_{a=30}^{50} f(a|07)\bar{Y}_{07,a}}.$$

For the year 2008, then, we have two cohorts of treated mothers:

$$d\bar{P}_{08} = \frac{\sum_{a=30}^{31} f(a|08) dP_{a-30} \bar{Y}_{08,a}}{\sum_{a=30}^{50} f(a|08)\bar{Y}_{08,a}}.$$

More generally, for $y = 07, \ldots, 28$, we have:

$$d\bar{P}_y = \frac{\sum_{a=30}^{30+y-07} f(a|y) dP_{a-30} \bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|y)\bar{Y}_{y,a}}.$$

Note that $d\bar{P}_y$ reaches its maximum in 2028 since then even the 50 year old mothers have been treated. Put differently: In the year 2028 and later, even those mother who gave birth in the reform year 2007 turn 50 and hence reach the upper bound of our sample age.

We now turn to the question how much of the full dynamic effect $d\bar{P}_{28}$ has been realized in an earlier year y, i.e., we are interested in the share $d\bar{P}_y/d\bar{P}_{28}$ for $y = 07, \ldots, 27$. It is given by:

$$\frac{d\bar{P}_y}{d\bar{P}_{28}} = \frac{\sum_{a=30}^{50} f(a|28)\bar{Y}_{28,a}}{\sum_{a=30}^{50} f(a|y)\bar{Y}_{y,a}} \cdot \frac{\sum_{a=30}^{30+y-07} f(a|y) dP_{a-30} \bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|28) dP_{a-30} \bar{Y}_{28,a}}. \qquad (*)$$

To gain an intuitive understanding of $(*)$, we now simplify it further by making more assumptions:

Assumption 2: *Calendar year effects are distributionally neutral in the sense that they scale up earnings for all age groups at the same rate. Denote by* $\rho_{07}^{y} = \frac{\bar{Y}_{28,a}}{\bar{Y}_{y,a}}$ *the relative increase in earnings for all age groups from year* $y = 07, ..., 27$ *to year* $y = 28$.

Under this assumption, our object of interest simplifies to

$$\frac{d\bar{P}_y}{d\bar{P}_{28}} = \frac{\sum_{a=30}^{50} f(a|28)\bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|y)\bar{Y}_{y,a}} \frac{\sum_{a=30}^{30+y-07} f(a|y) dP_{a-30}\bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|28) dP_{a-30}\bar{Y}_{y,a}}. \tag{$**$}$$

Assumption 3: *The change in the child penalty* $dP_t = dP$ *is independent of* t.

This assumption seems very strong at first sight. It is not only of pedagogical use, however, but also pretty close to the evidence presented in Figure 8.

Under this assumption, the equation further simplifies to

$$\frac{d\bar{P}_y}{d\bar{P}_{28}} = \frac{\sum_{a=30}^{30+y-07} f(a|y)\bar{Y}_{y,a}}{\sum_{a=30}^{50} f(a|y)\bar{Y}_{y,a}}. \tag{$***$}$$

Now consider the year 2018 ($y = 18$), which is our particular case of interest:

$$\frac{d\bar{P}_{18}}{d\bar{P}_{28}} = \frac{\sum_{a=30}^{41} f(a|18)\bar{Y}_{18,a}}{\sum_{a=30}^{50} f(a|18)\bar{Y}_{18,a}}.$$

In the numerator, we have the aggregated earnings of all mothers aged $30-41$ (i.e., all those that were treated in calendar year 2018 because they got their first child in 2007 or later). In the denominator, we have aggregated earnings of all mothers aged $30-50$. The ratio captures the share of aggregate maternal earnings that accrues to mothers who have been treated. The larger this share, the more of the long-run effect that the reform has on the average child penalty has materialized in calendar year 2018. To gain further intuition, we impose two more simplifying assumptions.

Assumption 5: *Assume that the age distribution is uniform with* $f(a|y) = 1/21$ *for all* a *and for all years* y.

Assumption 6: *Assume that potential earnings are independent of age* a, $\bar{Y}_{y,a} = \bar{Y}_{y,a'}$.

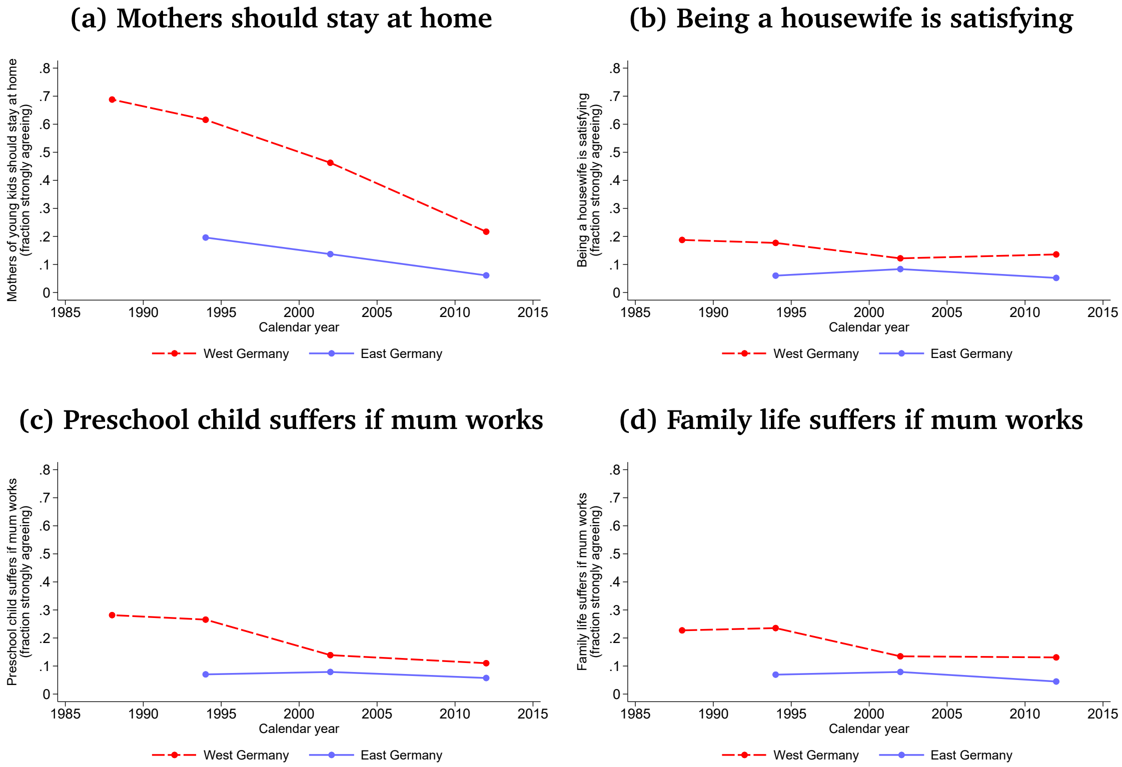
In this case, we obtain

$$\frac{d\bar{P}_{18}}{d\bar{P}_{28}} = \frac{18-07}{28-07} = \frac{11}{21}.$$

i.e., then 11/21 of the long-run effect have been realized. Clearly, Assumption 6 is restrictive because earnings generally increase with experience. Taking these life-cycle effects into account, the denominator of $d\bar{P}_{18}/d\bar{P}_{28}$ increases relative to the numerator. The implication is that – assuming the other assumptions still hold – less than 11/21 of the dynamic effect has realized in year 2018.

D Further figures

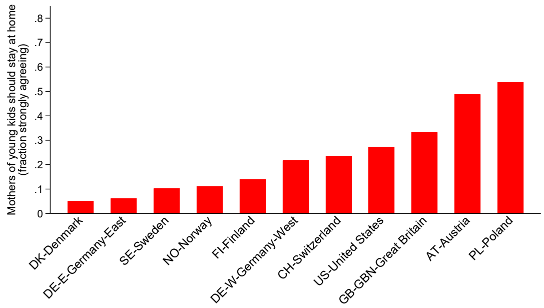
Figure A.1: Evolution of gender identity norms in Germany



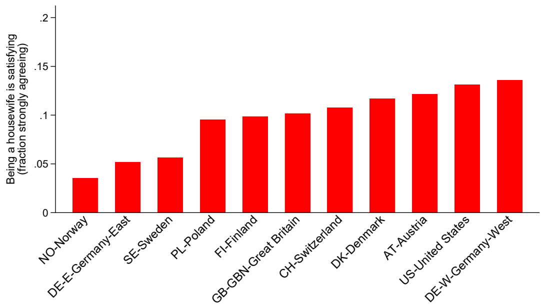
Notes: This figure demonstrates how gender identity norms developed in Germany over time. It relies on data from the International Social Survey Programme (waves: 1988, 1994, 2002, 2012). The exact survey questions read as follows: Figure A.1a: "Do you think that women should work outside the home full-time, part-time or not at all under the following circumstances? When there is a child under school age." The figure plots the fraction of respondents stating that a woman should stay at home. Figures A.1b to Figure A.1d, instead, show the fraction of respondents who "strongly agree" with the following statements: Figure A.1b: "Being a housewife is just as fulfilling as working for pay." Figure A.1c: "A preschool child is likely to suffer if his or her mother works." Figure A.1d: "All in all, family life suffers when the woman has a full-time job."

Figure A.2: International comparison of gender identity norms

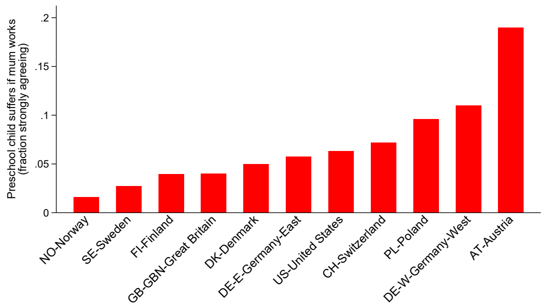
(a) Mothers should stay at home



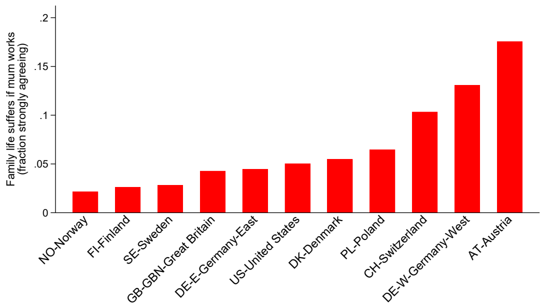
(b) Being a housewife is satisfying



(c) Preschool child suffers if mum works



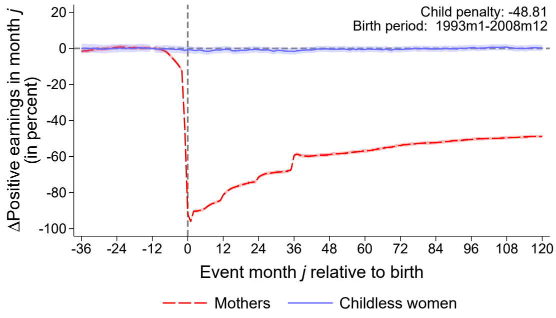
(d) Family life suffers if mum works



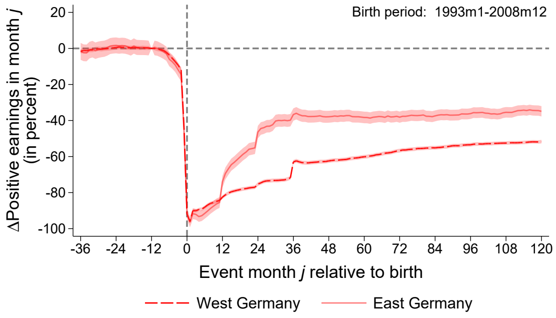
Notes: This figure compares gender identity norms across various countries. It relies on data from the International Social Survey Programme (wave: 2012). The exact survey questions read as follows: Figure A.2a: "Do you think that women should work outside the home full-time, part-time or not at all under the following circumstances? When there is a child under school age." The figure plots the fraction of respondents stating that a woman should stay at home. Figures A.2b to Figure A.2d, instead, show the fraction of respondents who "strongly agree" with the following statements: Figure A.1b: "Being a housewife is just as fulfilling as working for pay." Figure A.2c: "A preschool child is likely to suffer if his or her mother works." Figure A.2d: "All in all, family life suffers when the woman has a full-time job."

Figure A.3: Extensive margin child penalties for births between 1993 and 2008

(a) Entire Germany



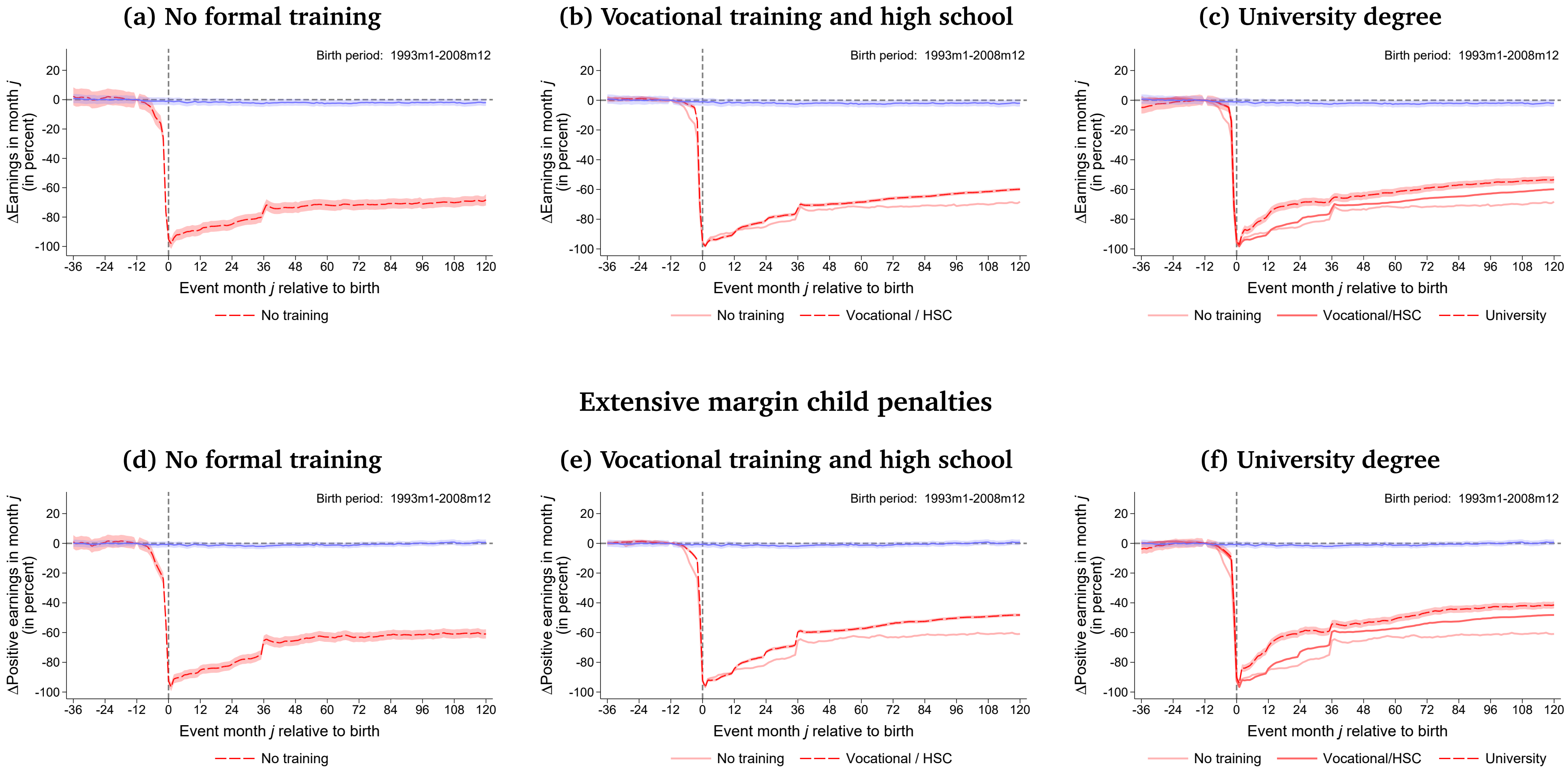
(b) West Germany versus East Germany



Notes: This figure presents the estimated extensive margin child penalties in mothers' earnings, $\widehat{P}_t$. In Figure A.3a, the dashed red line shows the child penalties for the entire country of Germany (including the Western and Eastern regions). The solid blue line represents the corresponding placebo effects for childless women. In the upper right corner of each figure, we report the extensive margin child penalties at event time 120 (i.e., ten years after birth). Figure A.3b depicts child penalties in East and West Germany separately. Both figures cover first births between 1993 and 2008 and consider event times from 36 months before to 120 months after the birth of the first child. The figures are based on regression that use a dummy indicating positive earnings as an outcome variable. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.4: Heterogeneity in child penalties by education

Child penalties in earnings

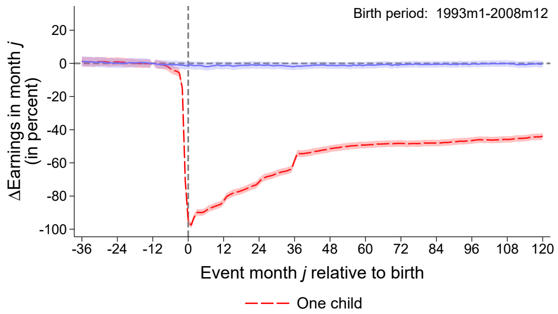


Notes: This figure shows how the child penalties around the birth of the first child, $\widehat{P}_t$, vary by education (no formal training, vocational training or high school diploma, and University degree). The dashed red lines represent the child penalties for the entire country of Germany (including the Western and Eastern regions). The solid blue line represents the corresponding placebo effects for childless women. The first row focuses on earnings as the outcome variable and the second row considers a dummy indicating positive earnings as an outcome. Both figures cover first births between 1993 and 2008 and consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

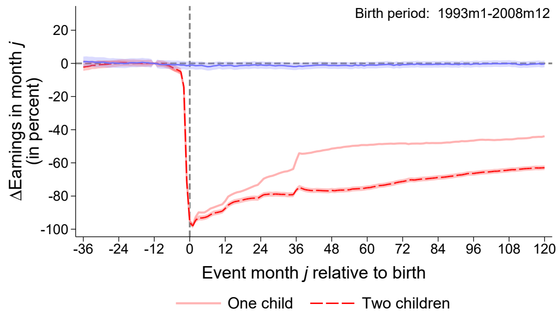
Figure A.5: Heterogeneity in child-related gaps by the number of children

Child penalties in earnings

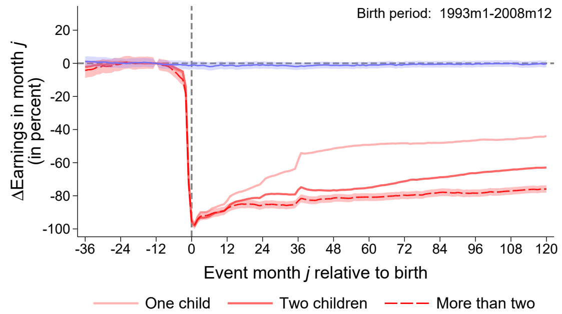
(a) One child



(b) Two children

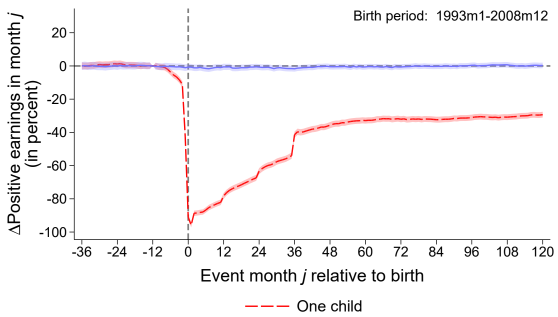


(c) More than two children

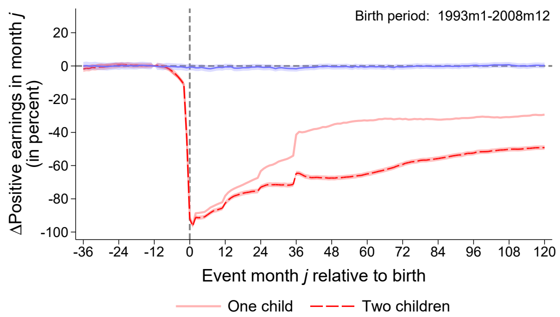


Extensive margin child penalties

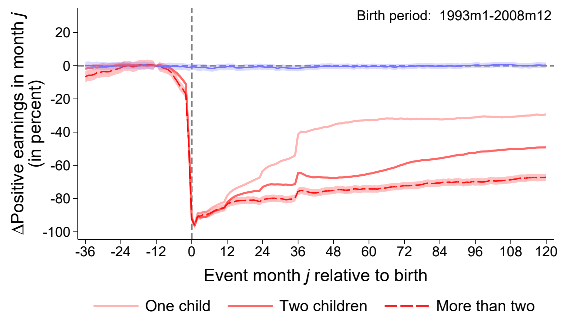
(d) One child



(e) Two children



(f) More than two children

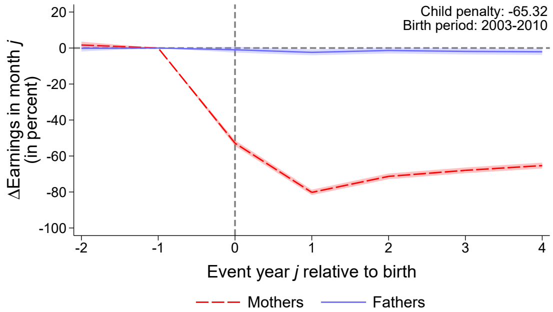


Notes: This figure shows how the child penalties around the birth of the first child, $\widehat{P}_t$, vary by the total number of children (one, two, and three or more). The dashed red lines represent the child penalties for the entire country of Germany (including the Western and Eastern regions). The solid blue line represents the corresponding placebo effects for childless women. The first row focuses on earnings as the outcome variable and the second row considers a dummy indicating positive earnings as an outcome. Both figures cover first births between 1993 and 2008 and consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

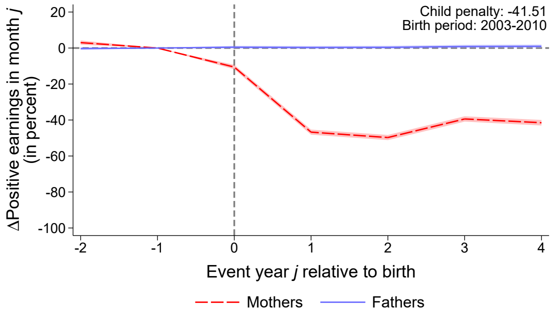
Figure A.6: Taxpayer Panel Data: Impacts of the first child

West Germany

(a) Child penalties in earnings

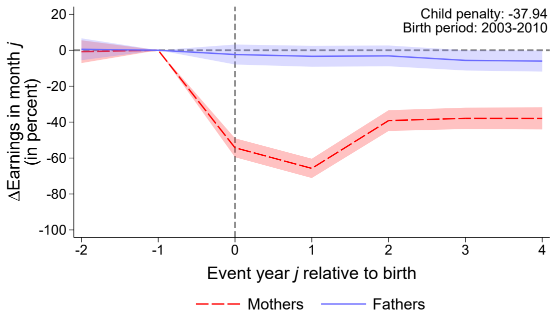


(b) Extensive margin child penalties

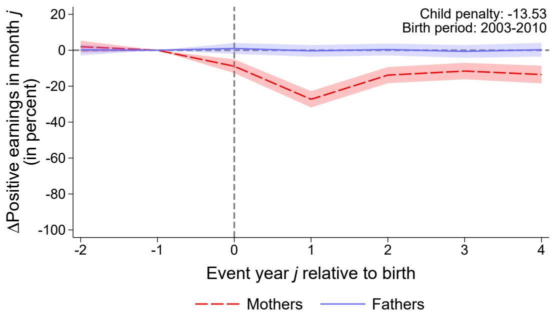


East Germany

(c) Child penalties in earnings

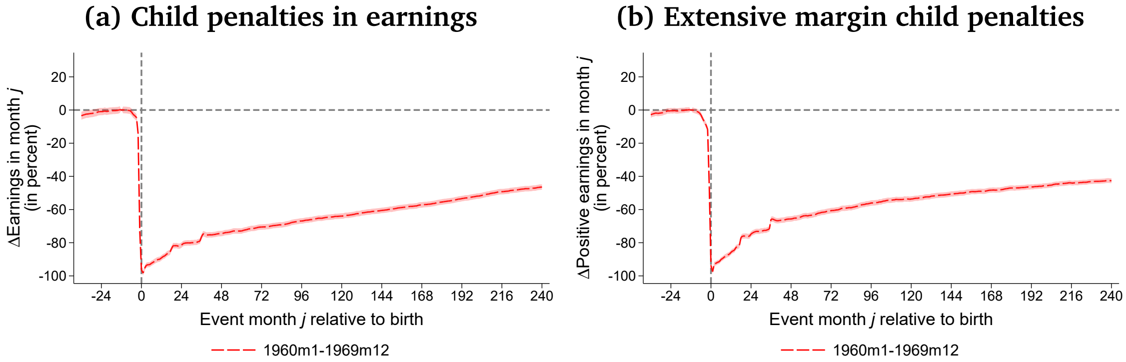


(d) Extensive margin child penalties



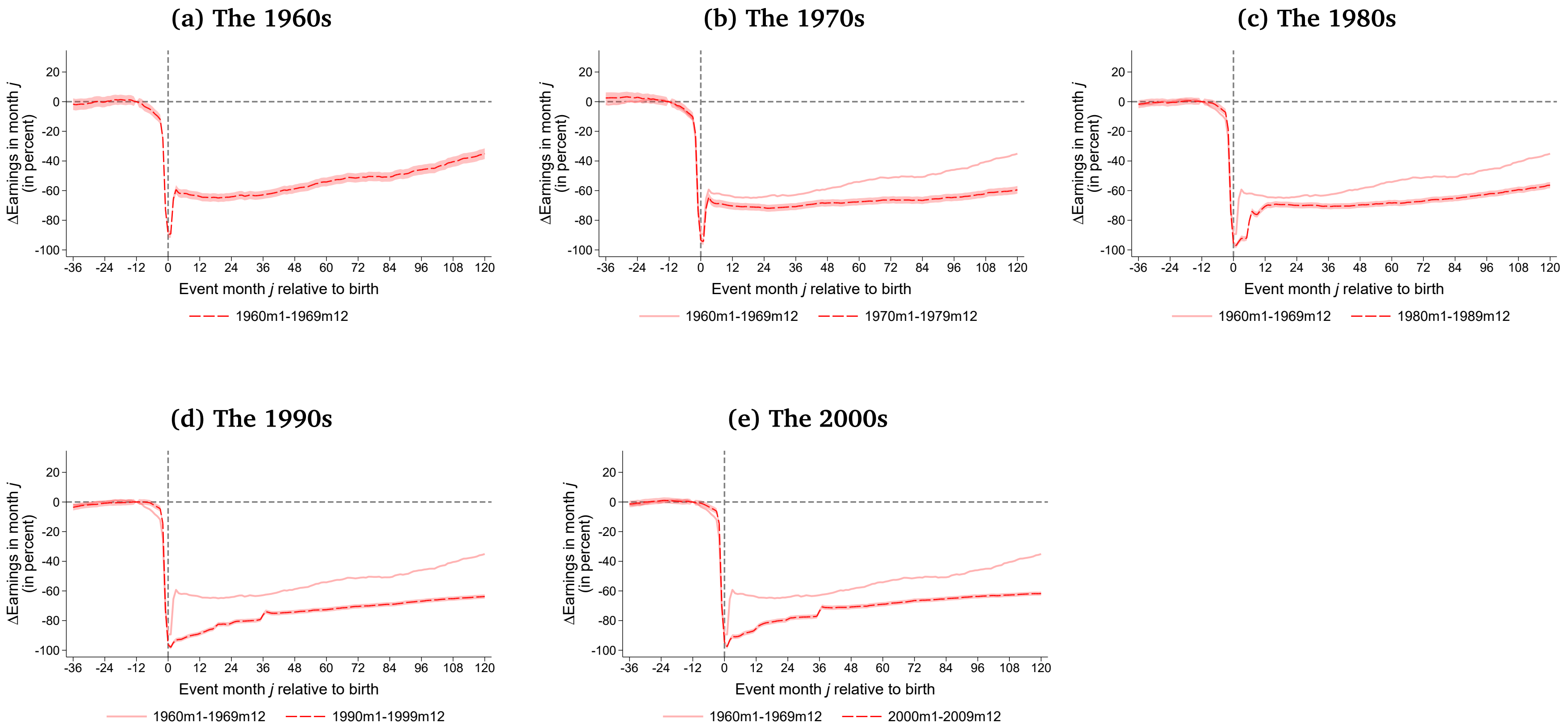
Notes: This figure presents the estimated child penalties for mothers (dashed red lines) and fathers (solid blue lines). In the upper right corner of each figure, we report the child penalties for the fourth year after birth. The German Taxpayer Panel serves as a data source (as it allows us to match children to both parents). The first row focuses on West Germany, and the second one on East Germany. Moreover, the left figures focus on earnings as the outcome variable and the right figures considers a dummy indicating positive earnings as an outcome. Both figures cover first births between 2003 and 2010 and consider event times from 2 years before to 4 years after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.7: Child penalties in the very long run



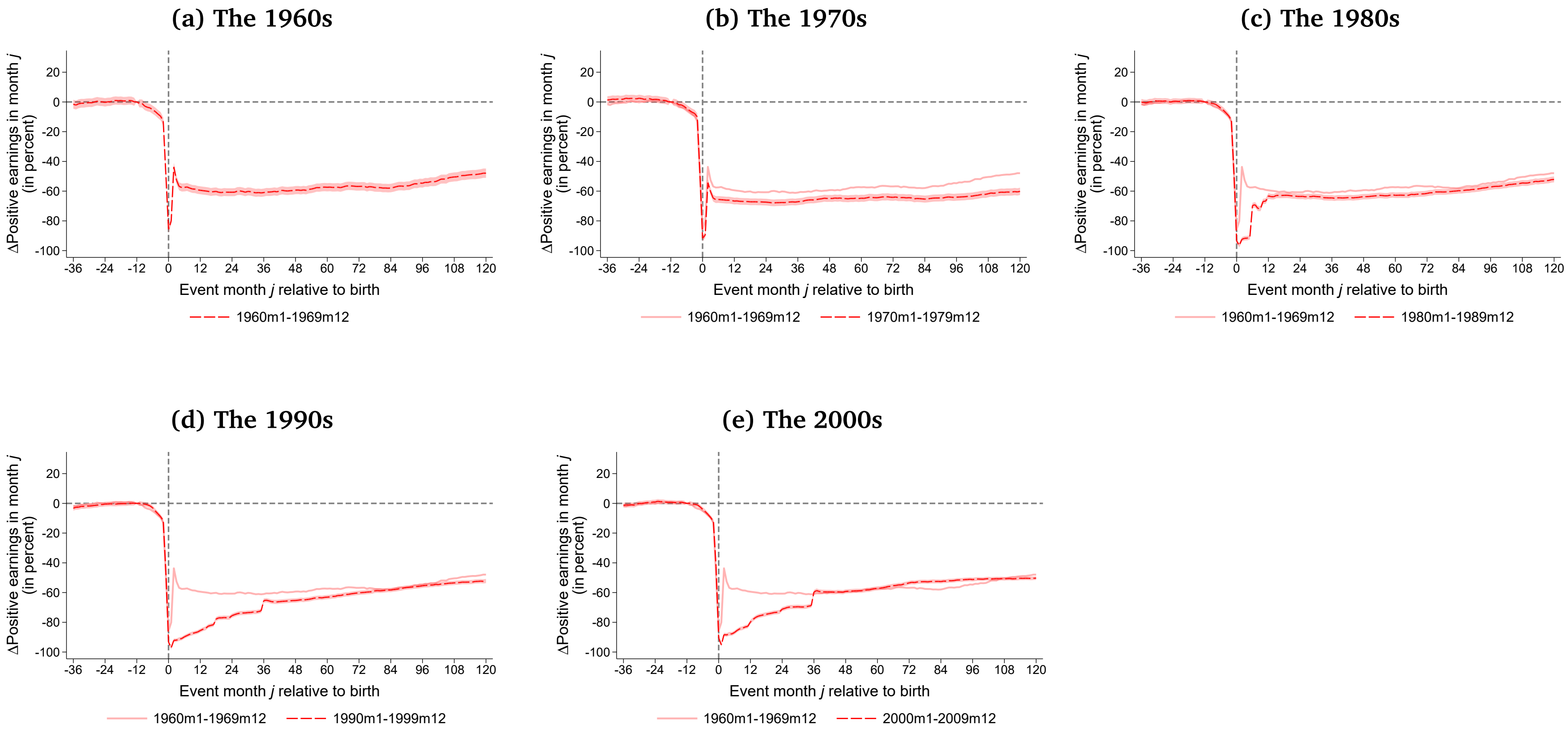
Notes: This figure shows the child penalties around the birth of the first child for West Germany, $\widehat{P}_t$, in the very long run (i.e., over 20 years). Figure A.7a focuses on earnings as the outcome variable and Figure Figure A.7b considers a dummy indicating positive earnings as an outcome. Both figures cover first births between 1990 and 1998 and consider event times from 36 months before to 240 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.8: Child penalties in earnings in West Germany in different decades



Notes: This figure shows how the child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). All figures consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

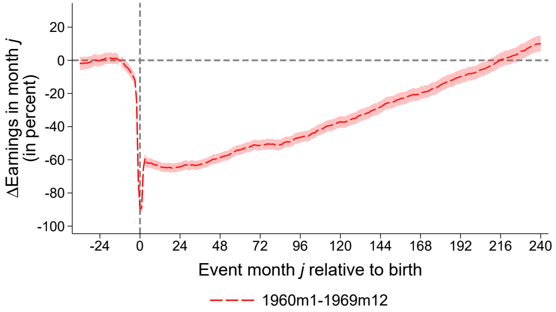
Figure A.9: Extensive margin child penalties in West Germany in different decades



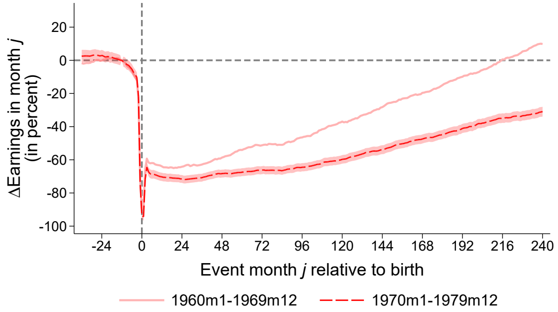
Notes: This figure shows how the extensive margin child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). All figures consider event times from 36 months before to 120 months after the birth of the first child. They rely on regressions that use a dummy indicating positive earnings as an outcome. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.10: Child penalties in earnings in West Germany in different decades: Very long run

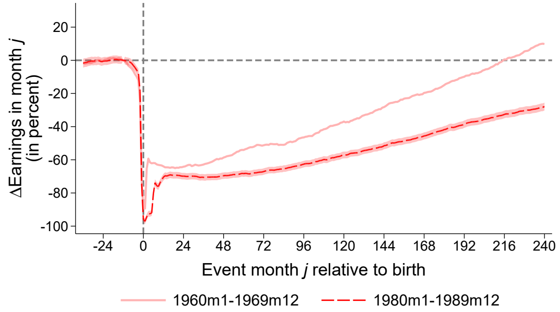
(a) The 1960s



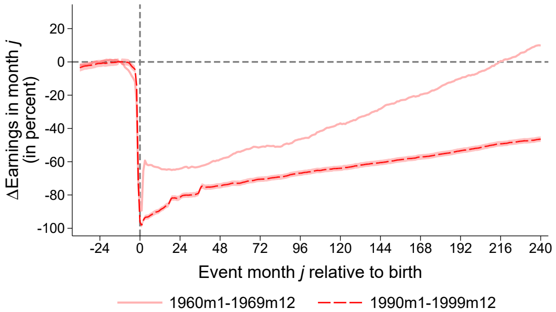
(b) The 1970s



(c) The 1980s



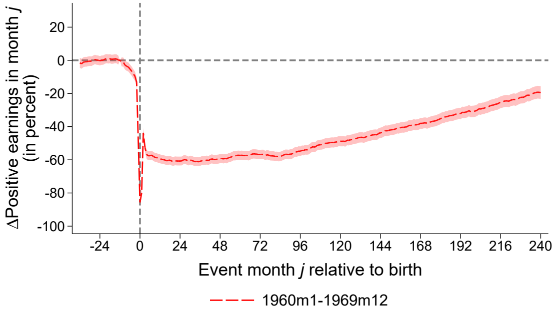
(d) The 1990s



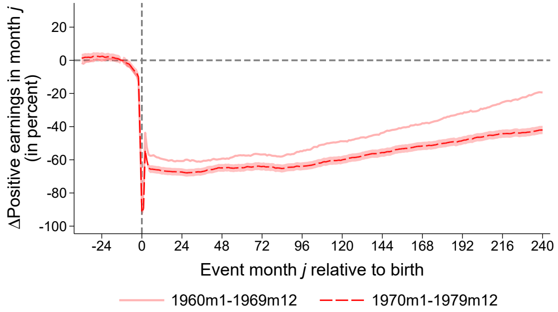
Notes: This figure shows how the very-long-run child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, and 1990s). All figures consider event times from 36 months before to 240 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.11: Extensive margin child penalties in West Germany in different decades: Very long run

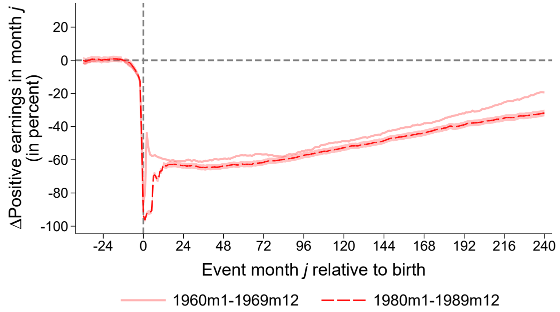
(a) The 1960s



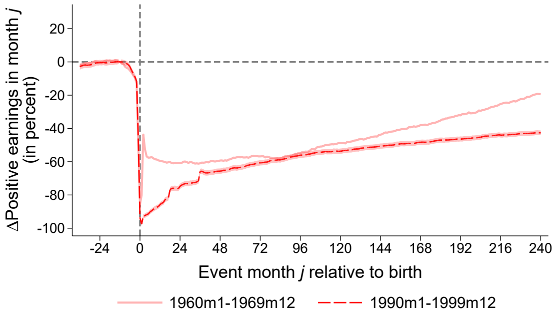
(b) The 1970s



(c) The 1980s

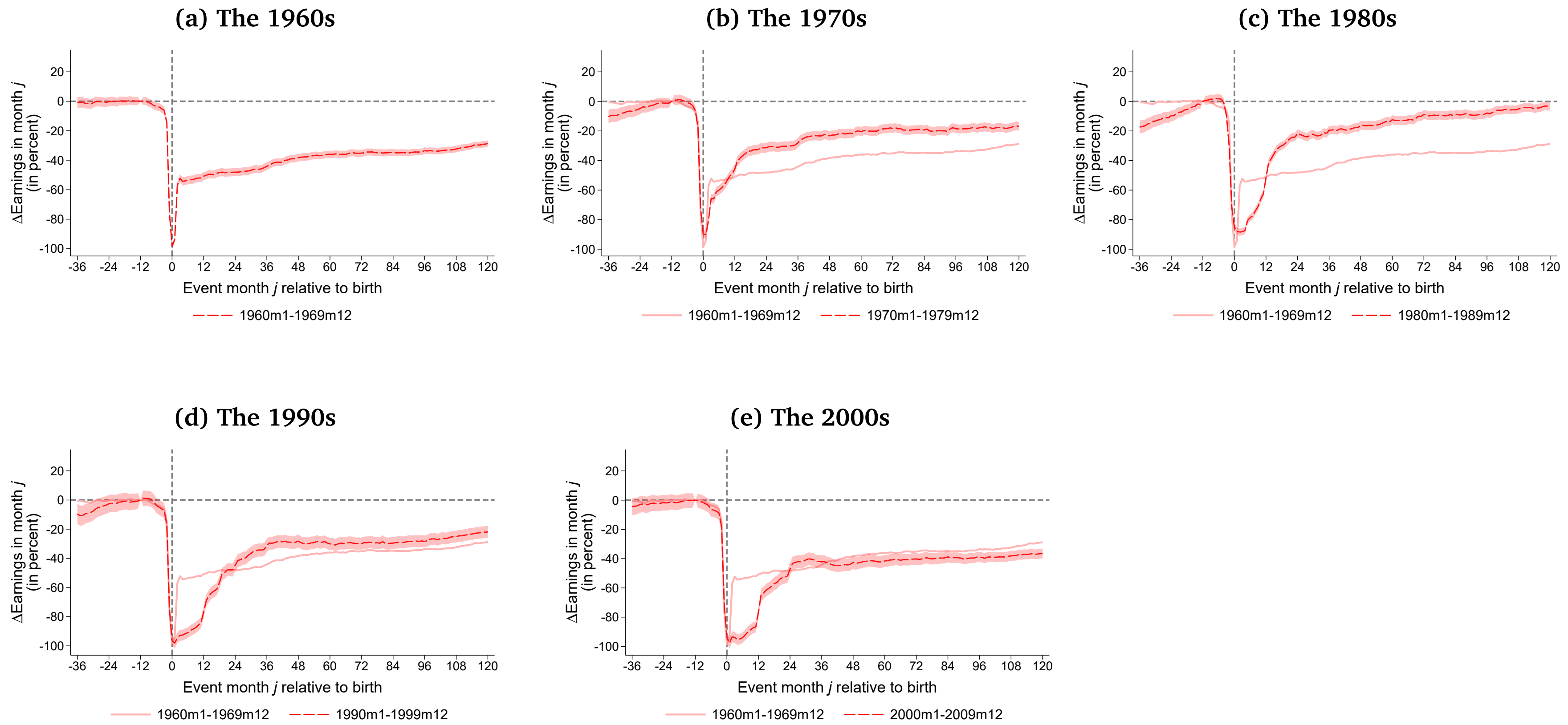


(d) The 1990s



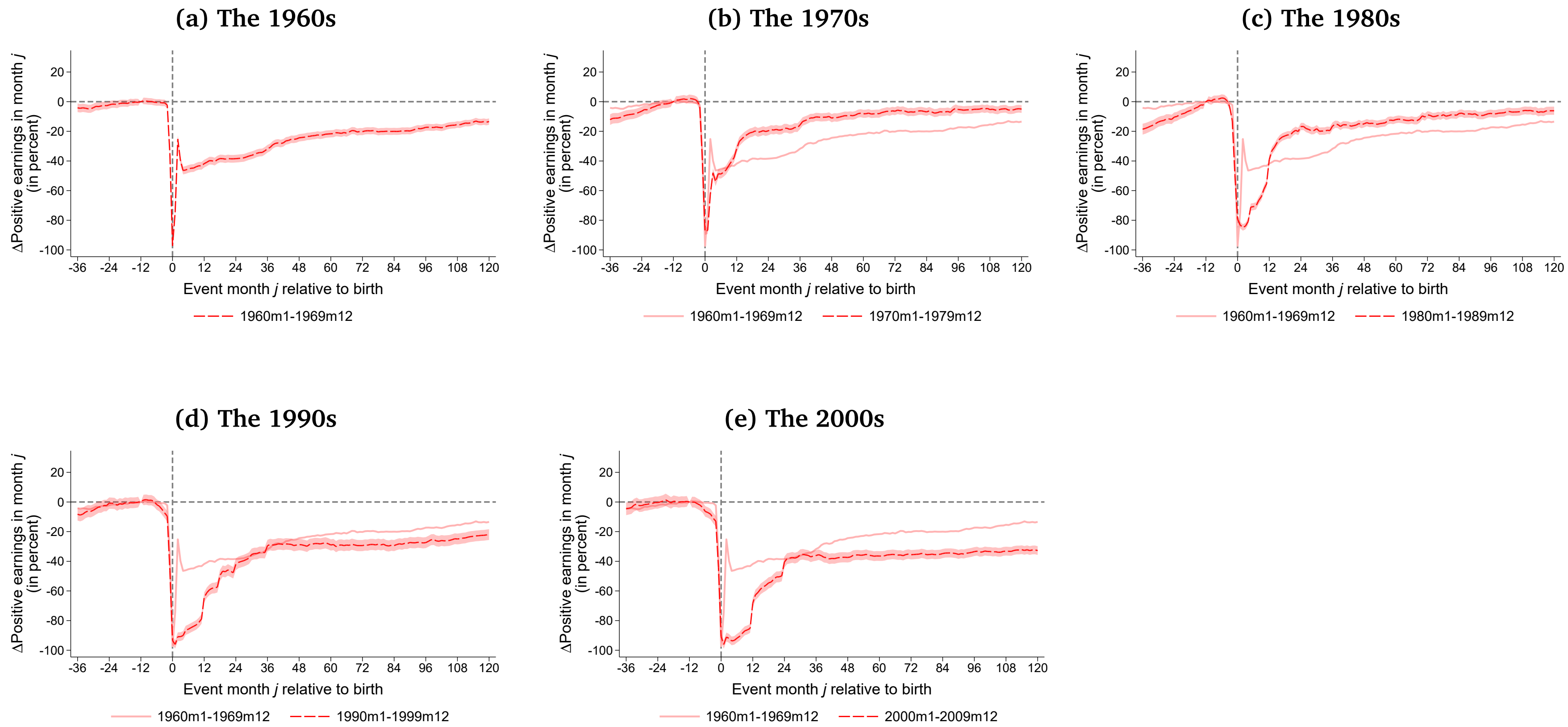
Notes: This figure shows how the very-long-run extensive margin child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). All figures consider event times from 36 months before to 120 months after the birth of the first child. They rely on regressions that use a dummy indicating positive earnings as an outcome. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.12: Child penalties in earnings in East Germany in different decades



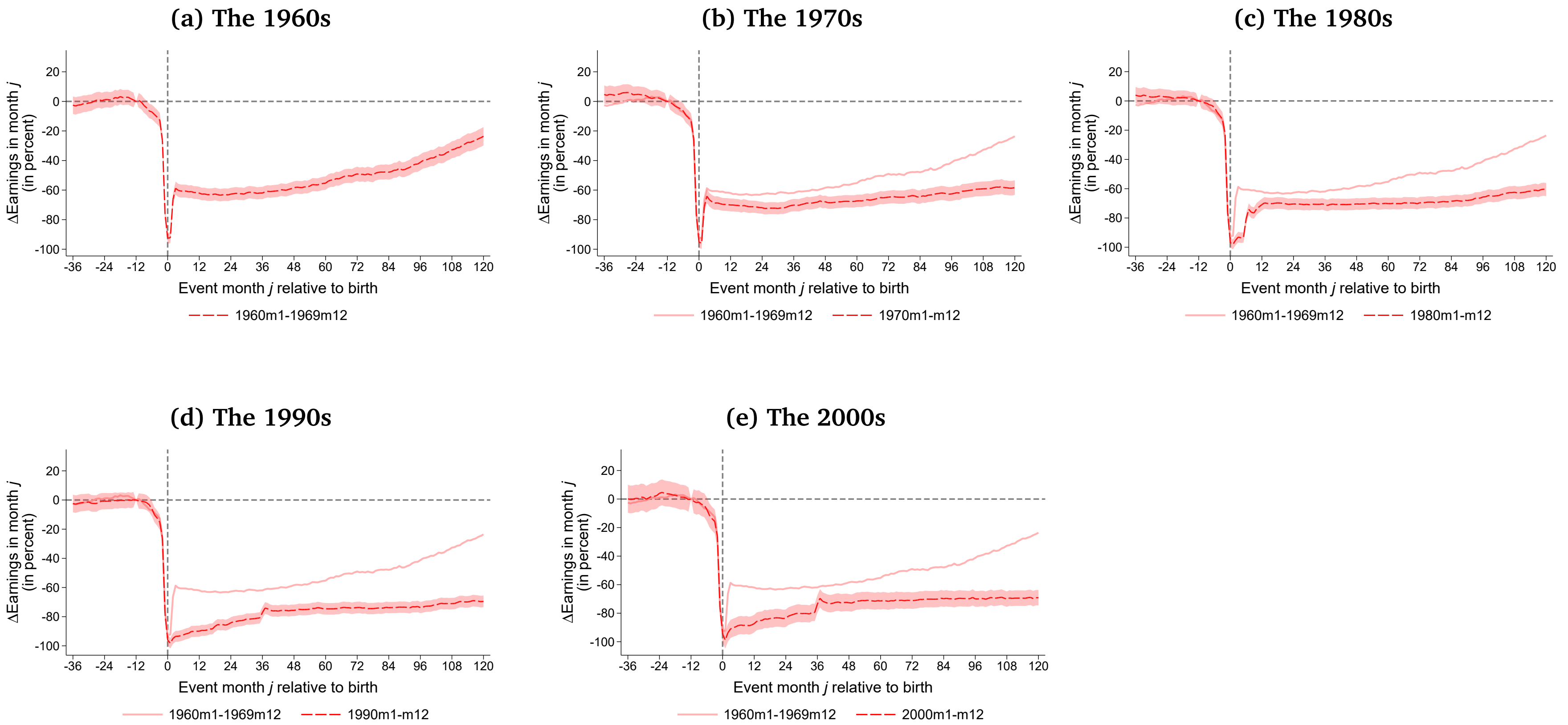
Notes: This figure shows how the child penalties in earnings in East Germany, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). All figures consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.13: Extensive margin child penalties in East Germany in different decades



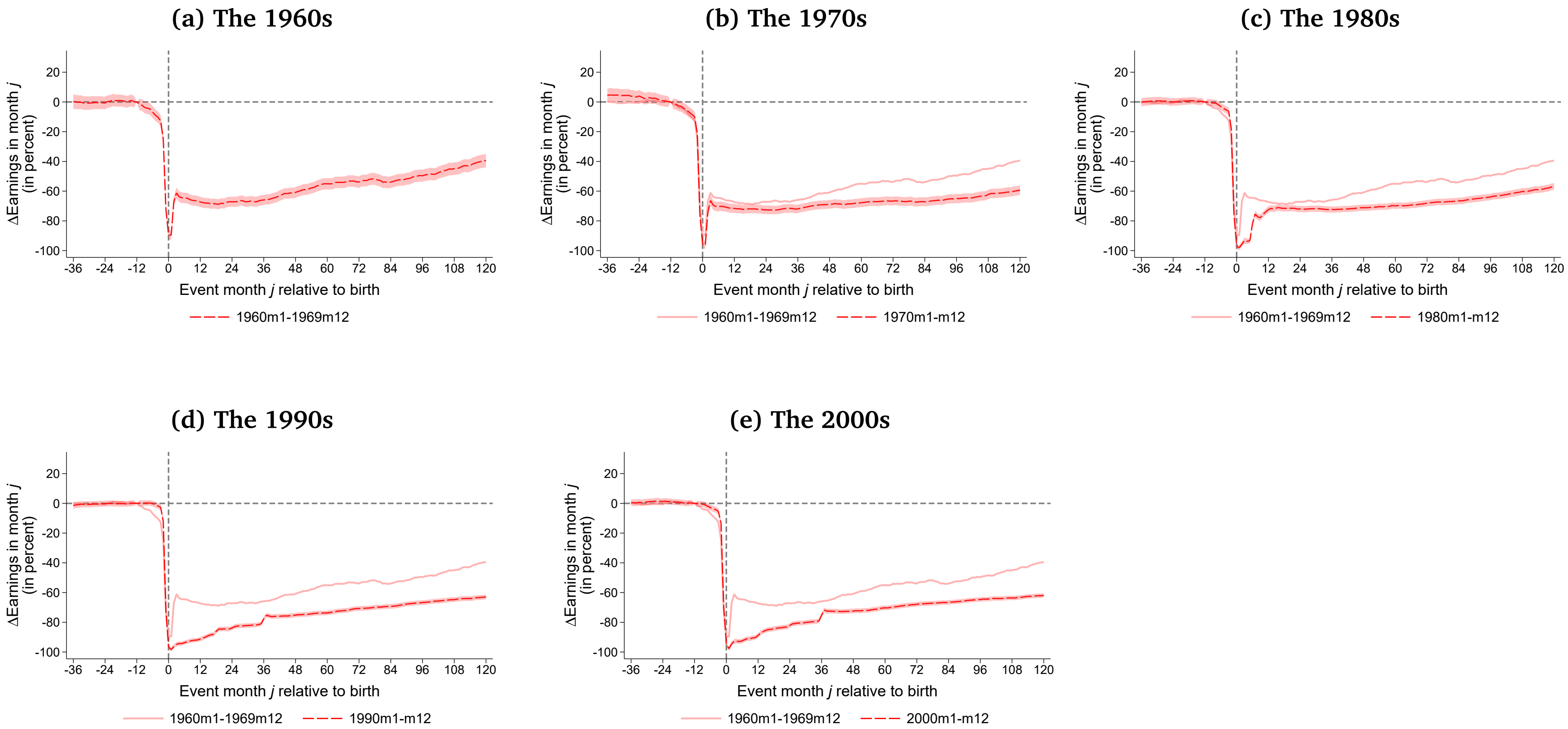
Notes: This figure shows how the extensive margin child penalties in earnings in East Germany, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). All figures consider event times from 36 months before to 120 months after the birth of the first child. They rely on regressions that use a dummy indicating positive earnings as an outcome. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.14: Child penalties in West Germany in different decades: No formal training



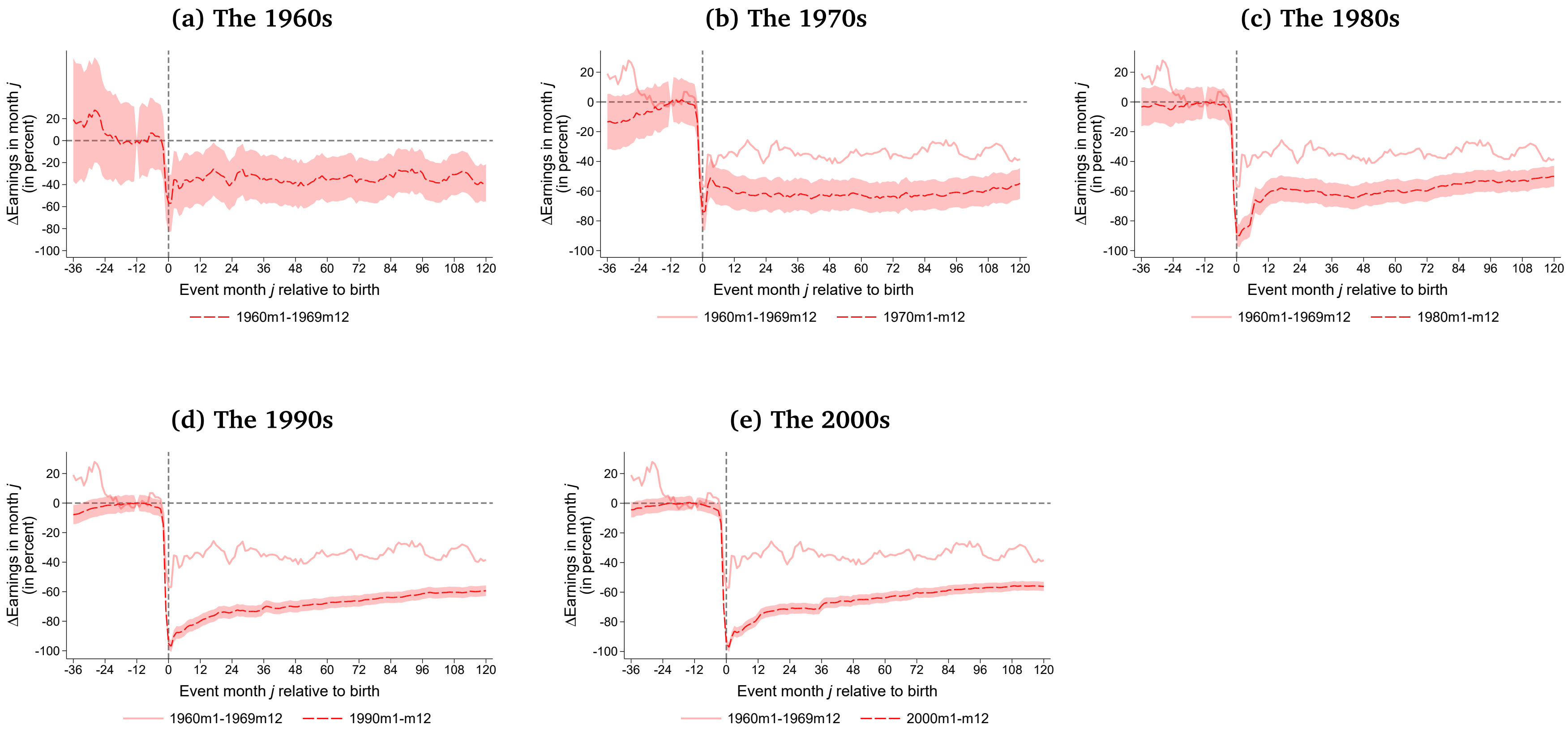
Notes: This figure shows how the child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). It focuses on mothers without any formal training. All figures consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.15: Child penalties in West Germany in different decades: Vocational training and high school



Notes: This figure shows how the child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). It focuses on mothers who either have vocational training or hold a high school degree. All figures consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

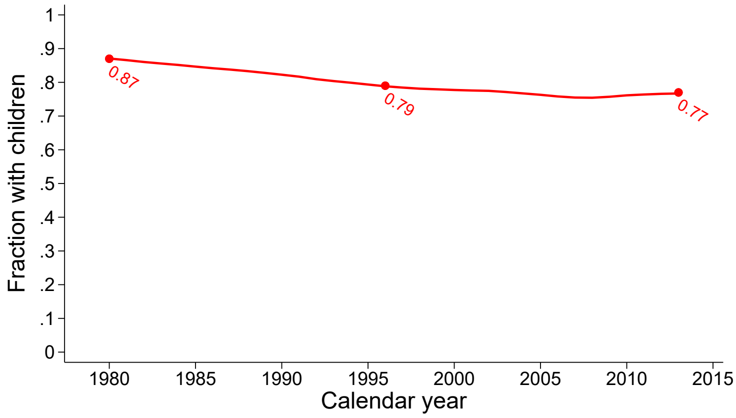
Figure A.16: Child penalties in West Germany in different decades: University degree



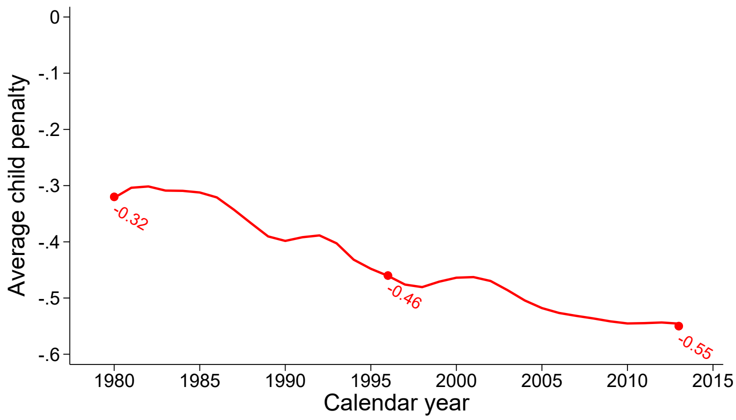
Notes: This figure shows how the child penalties in earnings, $\widehat{P}_t$, vary by the decade in which mothers gave birth to their first child (1960s, 1970s, 1980s, 1990s, and 2000s). It focuses on mothers without a University degree. All figures consider event times from 36 months before to 120 months after the birth of the first child. The shaded areas indicate heteroscedasticity-robust confidence intervals.

Figure A.17: Components of child-related inequality: Evolution over time

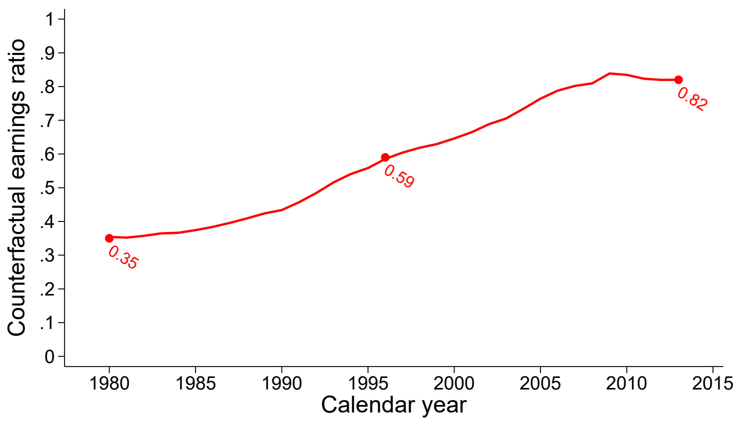
(a) Share of mothers in year y



(b) Average child penalty in year y



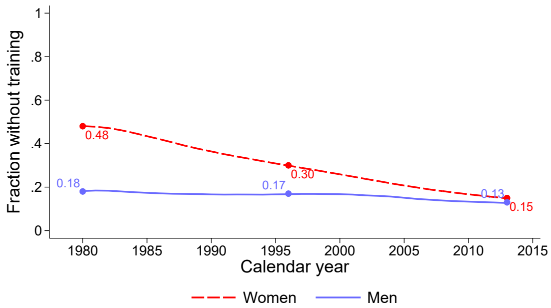
(c) Ratio of mothers' counterfactual earnings to male earnings



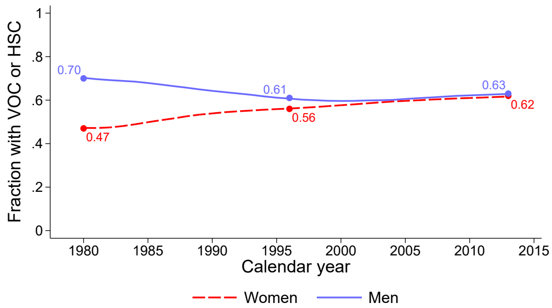
Notes: This figure depicts the evolution of the three components of child-related inequality Δ_y^r. Figure A.17a focuses on the share of women with children (ϕ_y), Figure A.17b shows the average child penalty ($\overline{P}_y$), and Figure A.19c demonstrates the counterfactual earnings ratio (Ψ_y). We compute each component for the sample of men and women aged 25 to 45 and focus on the years 1980 to 2013.

Figure A.18: Education: Evolution over time

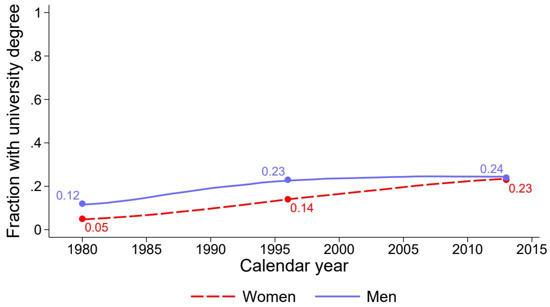
(a) No formal training



(b) Vocational training and high school



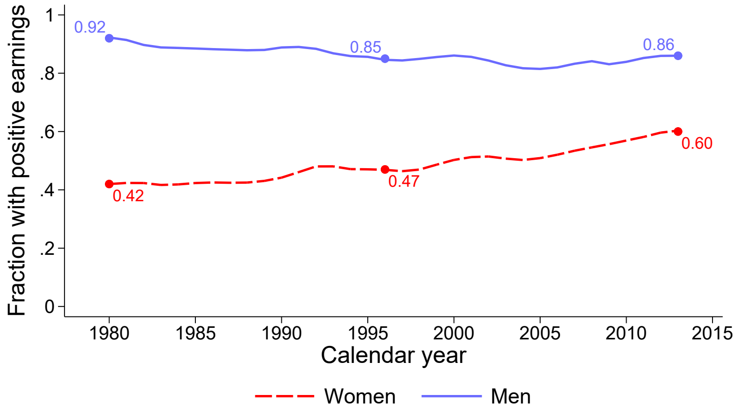
(c) University degree



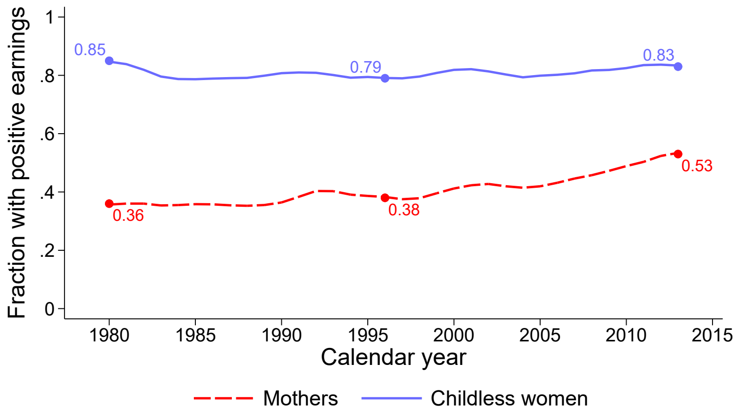
Notes: This figure shows how educational levels evolve over time. It shows the fraction of all women (dashed lines) and all men (solid lines) in year y without formal training (Figure A.18a), with vocational training or a high school degree (Figure A.18b), and a University degree (Figure A.18c). We compute those fractions for the sample of people aged 25 to 45 and focus on the years 1980 to 2013.

Figure A.19: Extensive margin of labor supply: Evolution over time

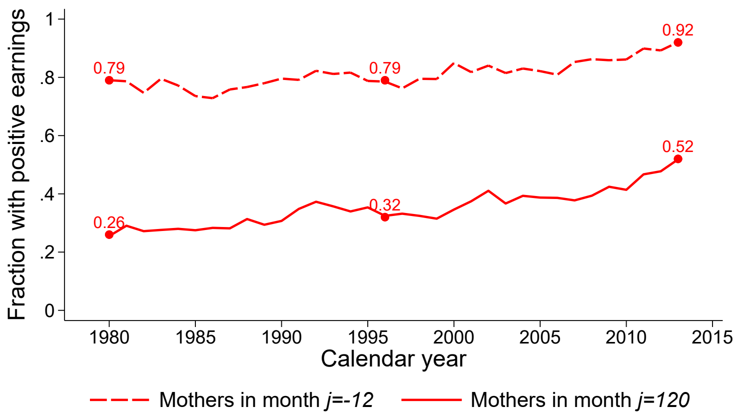
(a) Women vs. men



(b) Mothers vs. childless women



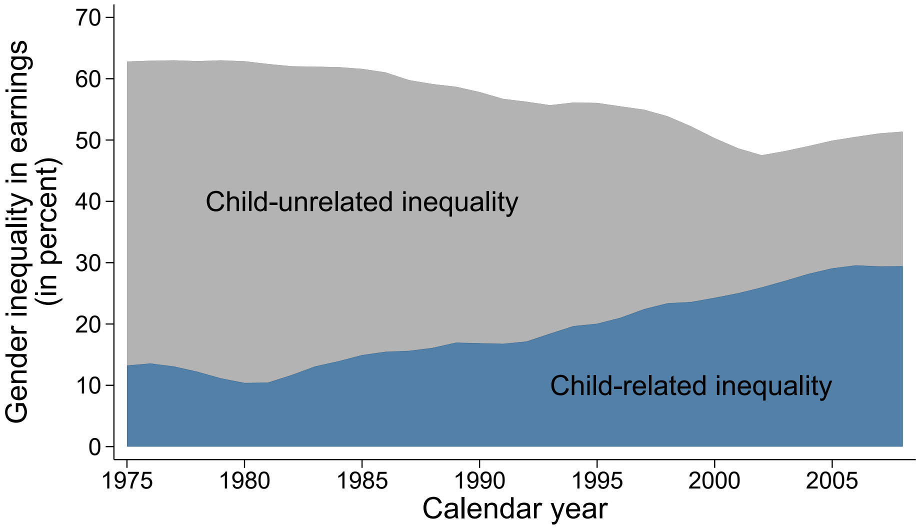
(c) Pre vs. post-birth behavior



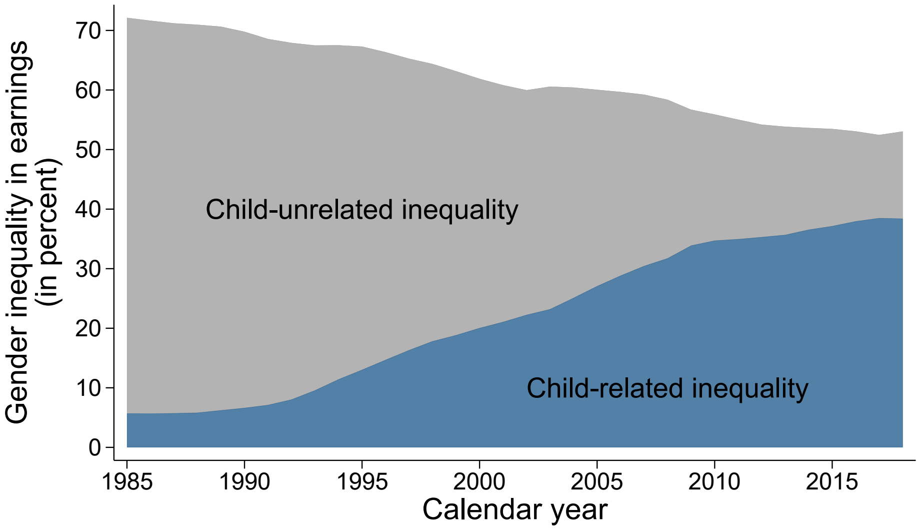
Notes: This figure depicts how the extensive margin of labor supply behavior evolves over time. It focuses on several subgroups: Figure A.19a shows the fraction of all women (dashed red line) and all men (solid blue line) who have positive earnings in year y. Figure A.19b, instead, depicts the fraction of all mothers (dashed red line) and all childless women (solid red line) who have positive earnings in year y. Lastly, Figure A.19a shows the fraction of all mothers in year y who have positive earnings at event time $j = -12$ (dashed red line) and event time $j = 120$ (solid red line). We compute those fractions for the sample of people aged 25 to 45 and focus on the years 1980 to 2013.

Figure A.20: Alternative decomposition samples

(a) Men and women between the ages of 20 and 40 (1975-2008)



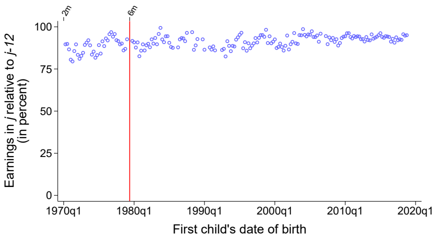
(b) Men and women between the ages of 30 and 50 (1985-2018)



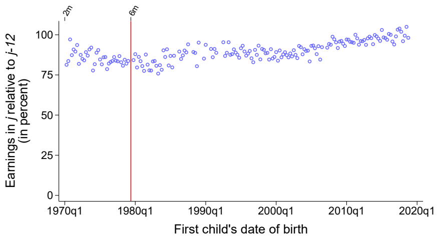
Notes: This figure decomposes the overall gender inequality in earnings into a child-related part and a child-unrelated residual part (see equation 4). It focuses on West Germany. The underlying event study models allow for year-y-specific event coefficients and control for year and age dummies. The child-related part reflects the effect of P_{yj}. By contrast, the child-unrelated part represents all the other drivers of the overall gap. The sample in Figure A.20a (Figure A.20b) spans the years from 1975 to 2008 (1985 to 2018) and includes men and women between the ages of 20 and 40 (30 and 50).

Figure A.21: Impacts of the 1979 and 1986 reforms on mothers' earnings in the raw data

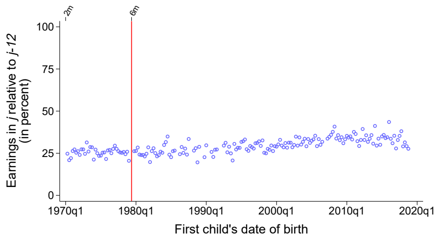
(a) 24 months before birth



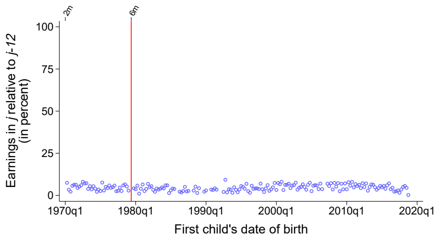
(b) 2 months before birth



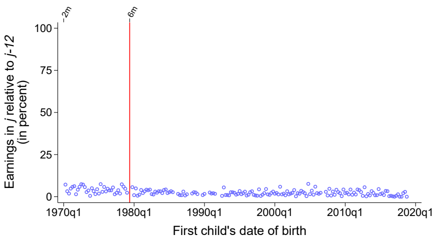
(c) 1 month before birth



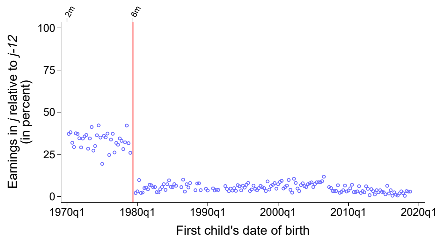
(d) Month of birth



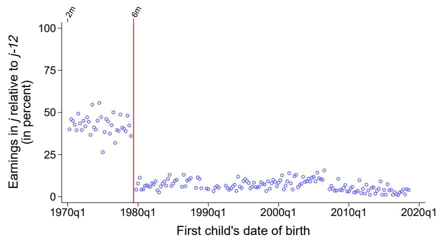
(e) 1 month after birth



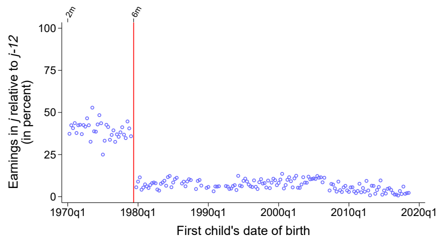
(f) 2 months after birth



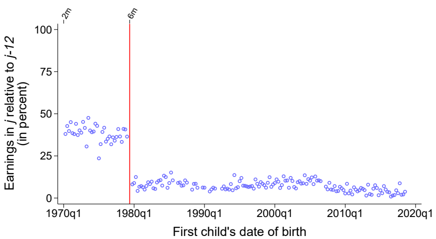
(g) 3 months after birth



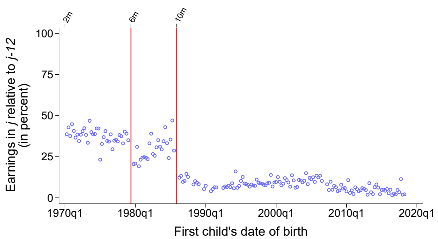
(h) 4 months after birth



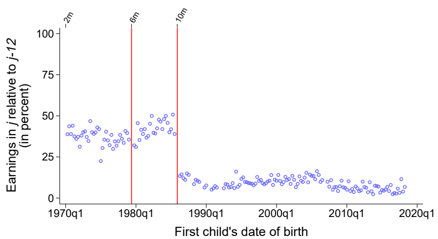
(i) 5 months after birth



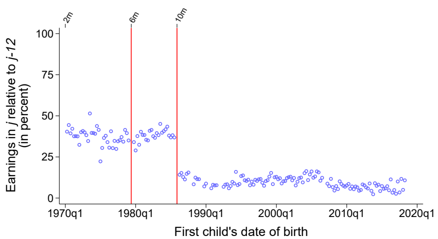
(j) 6 months after birth



(k) 7 months after birth



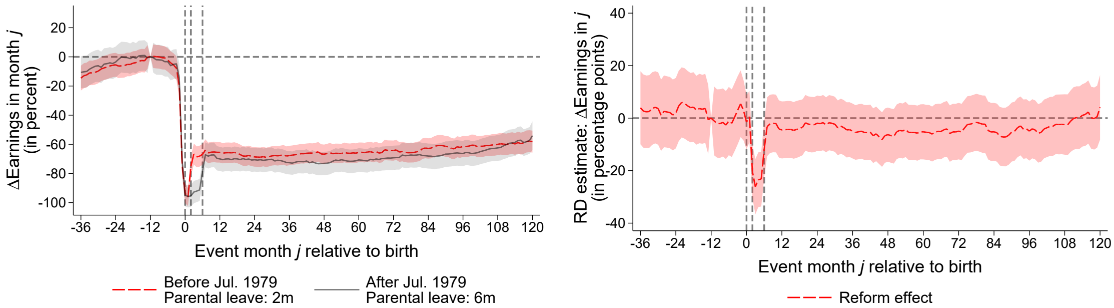
(l) 8 months after birth



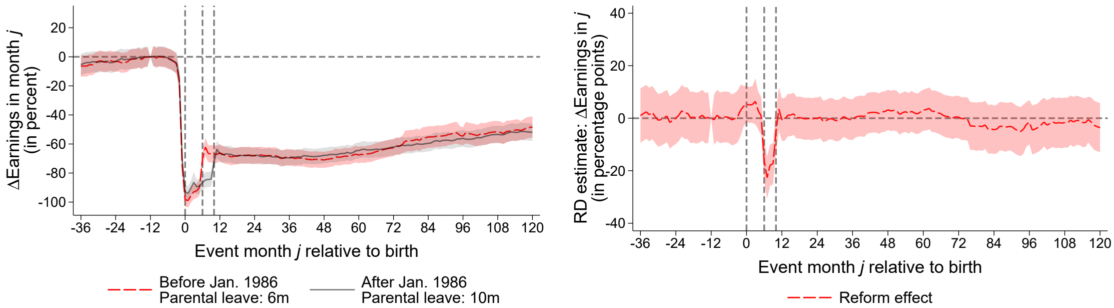
Notes: This figure plots the raw data to visualize the impacts of the 1979 and 1986 reforms descriptively. Each of the 12 figures focuses on a specific event time (i.e., on mothers at t months before/after birth). Specifically, for each event time, we display mothers' average earnings as a percentage of their earnings 12 months prior to birth, plotted against the month their first child was born. Four observations stand out. Each circle, thus, represents an average value for mothers who give birth in one particular month.

Figure A.22: Impacts of parental leave reforms

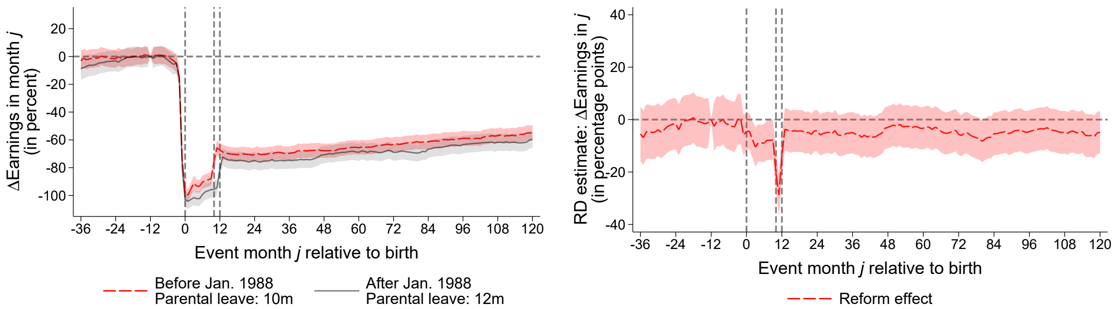
(a) Parental leave reform in 1979



(b) Parental leave reform in 1986



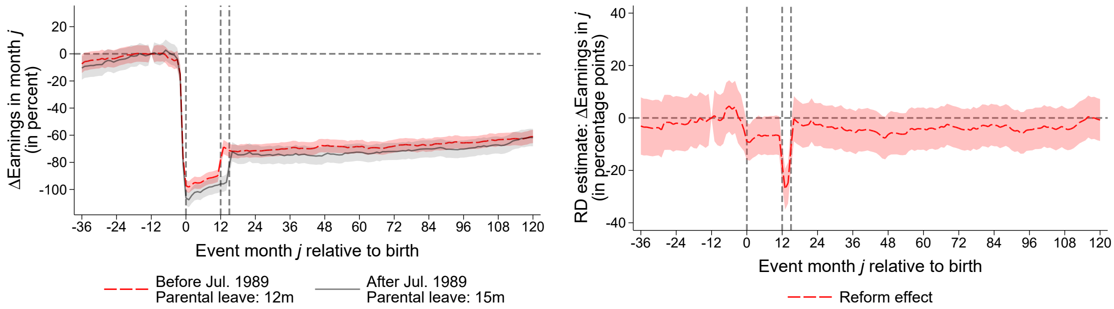
(c) Parental leave reform in 1988



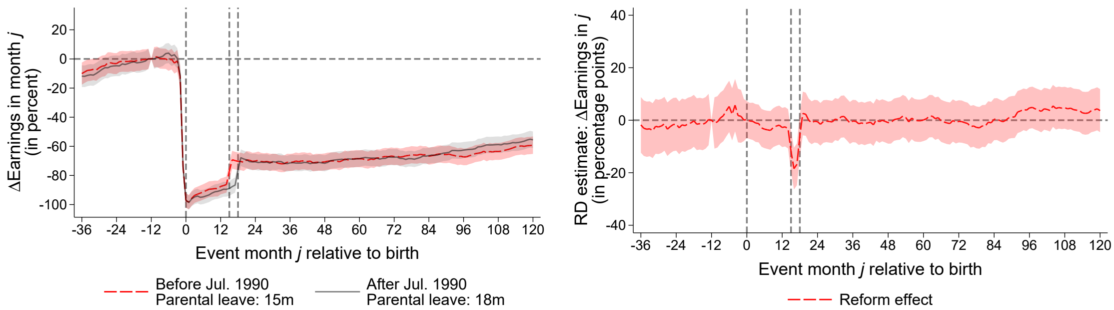
Notes: These figures visualize the impacts of paid parental leave reforms in 1979, 1986, and 1988 on mothers' earnings trajectories in West Germany. It focuses on first-time mothers. The left figures depict the percentage impacts of children on the earnings of mothers who are just treated (solid line) or untreated (dashed line) by the reform. By contrast, the right figures delineate the corresponding RD reform effects.

Figure A.23: Impacts of parental leave reforms

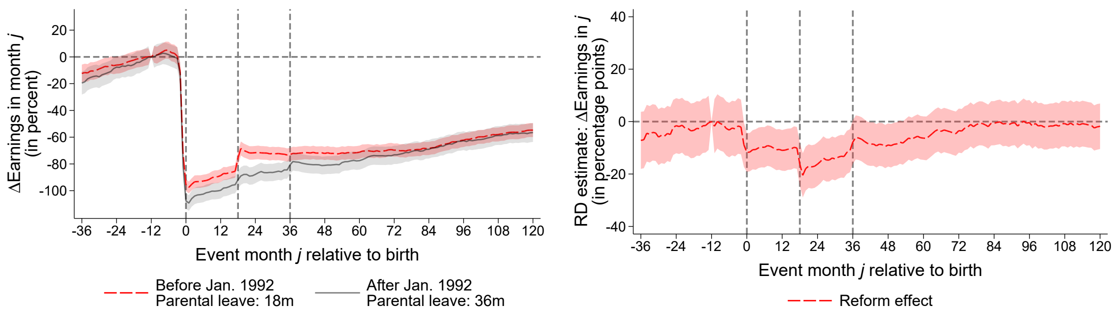
(a) Parental leave reform in 1989



(b) Parental leave reform in 1990

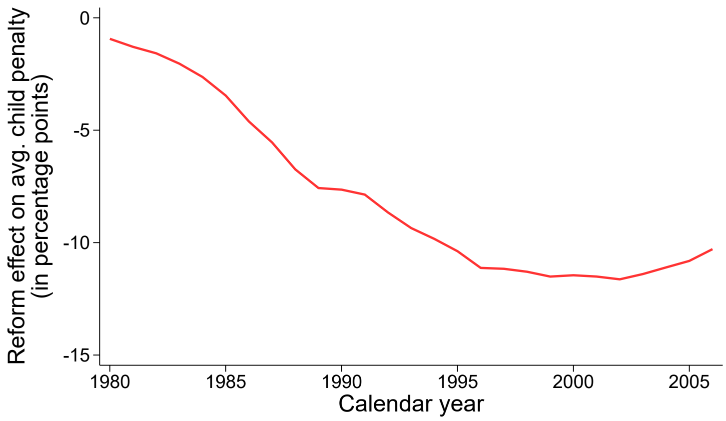


(c) Parental leave reform in 1992



Notes: These figures visualize the impacts of paid parental leave reforms in 1989, 1990, and 1992 on mothers' earnings trajectories in West Germany. It focuses on first-time mothers. The left figures depict the percentage impacts of children on the earnings of mothers who are just treated (solid line) or untreated (dashed line) by the reform. By contrast, the right figures delineate the corresponding RD reform effects.

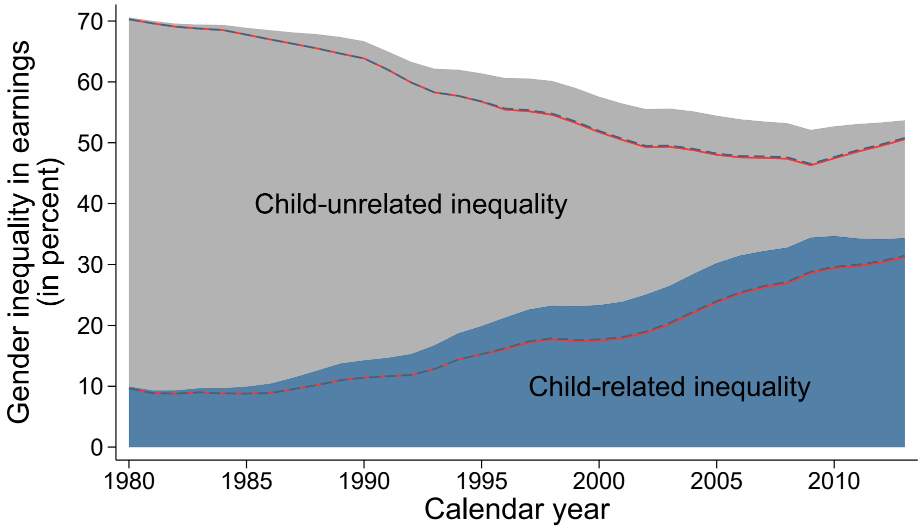
Figure A.24: Joint effect of all parental leave reforms on average child penalty



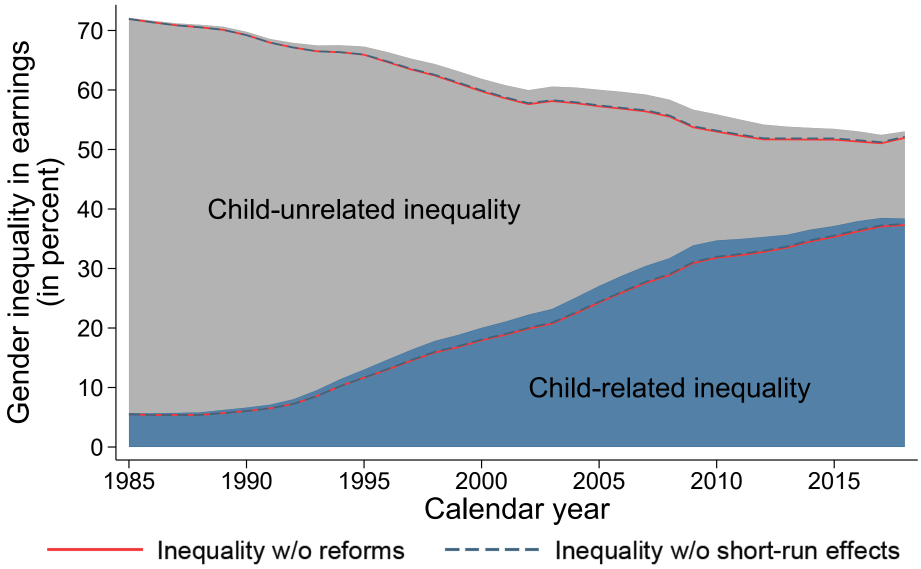
Notes: This figure depicts how the six parental leave reforms jointly affected the average child penalty $\bar{P}_y$ in year y. It only considers statistically significant effects. The computation of the joint effect follows equation (12). We calculate the joint effect for the sample of men and women aged 25 to 45 and focus on the years 1980 to 2006.

Figure A.25: Alternative samples: Impact of job-protected parental leave reforms

(a) Men and women between the ages of 25 and 45 (1980-2013)

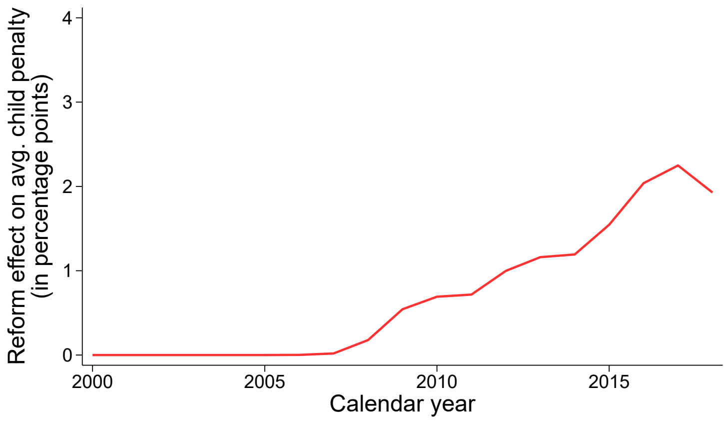


(b) Men and women between the ages of 30 and 50 (1985-2018)



Notes: This figure illustrates how the parental leave reforms between 1979 and 1992 affected overall gender inequality (blue plus gray) and child-related inequality (blue) in earnings in West Germany. We consider two counterfactual scenarios. Scenario 1 (solid red lines) depicts how overall inequality and child-related inequality would have evolved without the leave reforms, accounting for their long-run and short-run effects on child penalties. Scenario 2 (long-dashed blue lines) depicts how inequality was changed by the reforms' short-run effects only. The underlying event study models allow for cohort-specific event time coefficients and control for year and age dummies. The sample in Figure A.25a (Figure A.25b) spans the years from 1980 to 2013 (1985 to 2018) and includes men and women between the ages of 25 and 45 (30 and 50).

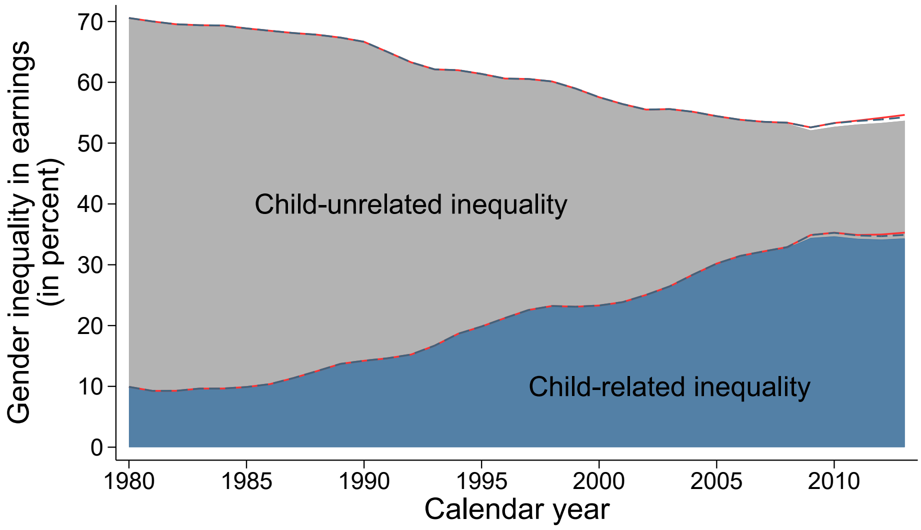
Figure A.26: Effect of parental benefit reform on average child penalty



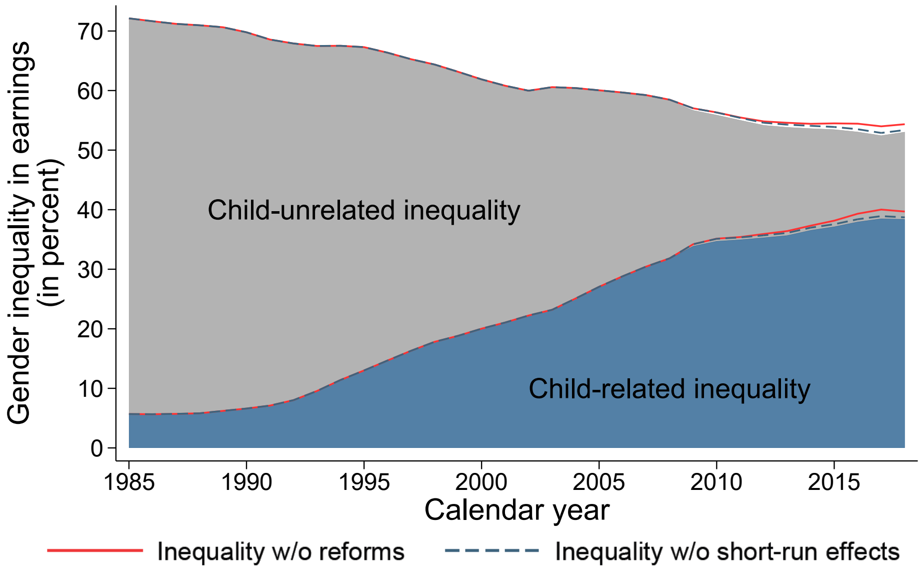
Notes: This figure depicts how the parental benefit reform affects the average child penalty $\bar{P}_y$ in year y. It only considers statistically significant effects. The computation follows equation (12). We calculate the joint effect for the sample of men and women aged 30 to 50 and focus on the years 2000 to 2018.

Figure A.27: Alternative samples: Impact of 2007 parental benefit reform

(a) Men and women between the ages of 25 and 45 (1980-2013)



(b) Men and women between the ages of 30 and 50 (1985-2018)



Notes: This figure illustrates how the parental benefit reform in 2007 affected overall gender inequality (blue plus gray) and child-related inequality (blue) in earnings in West Germany. We consider two counterfactual scenarios. Scenario 1 (solid red lines) depicts how overall inequality and child-related inequality would have evolved without the leave reforms, accounting for their long-run and short-run effects on child penalties. Scenario 2 (long-dashed blue lines) depicts how inequality was changed by the reforms' short-run effects only. The underlying event study models allow for cohort-specific event time coefficients and control for year and age dummies. The sample in Figure A.27a (Figure A.27b) spans the years from 1980 to 2008 (1985 to 2018) and includes men and women between the ages of 25 and 45 (30 and 50). Figure A.27b is equal to Figure 9 in the main text; it is included to simplify a comparison.