Female Labor Supply, Fertility and Parental Leave Policy Design

Mary Ann Bronson* Deniz Sanin†

July 17, 2024
Preliminary

Abstract

Parental leave (PL) policies vary significantly across countries, creating different incentives for women to work and have children. Existing empirical studies, however, find mostly negligible effects of PL reforms on these variables, leading to a growing consensus that PL design has little influence on women's decisions (e.g., Kleven et al. 2021). In this paper, we argue that existing empirical studies – which mostly exploit regression discontinuities around PL reform dates – offer exceptionally clean identification, but estimate parameters that are not easily interpretable and are potentially misleading to a policymaker. We proceed in three steps. First, we document the main differences between the ideal experiment to evaluate a PL reform and the typical quasi-experiment in the literature; we show that, by construction, the latter is likely to yield only negligible effects of PL policies, even if the true policy effect is large. Next, using event studies around first birth from the U.K., Sweden, Austria, and Germany, we provide descriptive evidence that the differences in women's work, PL take-up, and childbearing across those countries are strongly in line with the incentives created by their respective PL designs. Based on this evidence, we then build a dynamic model of women's labor supply and fertility, which can simulate a range of policy designs that differ in length, compensation structure, and eligibility rules. We show that the model replicates the negligible RDD estimates of past policy reforms, even as the true effect of different policy reforms is large, affecting medium- and long-run labor force participation rates by as much as much as 10 percentage points. We calibrate the model to U.K. data, and show that when we counterfactually expose U.K. women to Swedish, Austrian, or German policies, they make work and childbearing decisions similar to those of women in those countries, consistent with the strong incentives embedded in different in PL policies. Finally, we consider the policy designs that maximize women's labor force participation, fertility, and welfare, respectively.

Keywords: Female labor supply, fertility, parental leave policy, Europe
JEL Codes: J13, J18, J22

*Georgetown University, Department of Economics, mary.ann.bronson@georgetown.edu
†University of South Carolina, Department of Economics, deniz.sanin@moore.sc.edu

1 Introduction

Parental leave (PL) policies differ significantly across countries. Austria, for example, spends on average $7,000 per child in PL expenditures, while Sweden spends more than $30,000 (OECD, 2021). The duration of paid leave and of job protection, how compensation is determined, the generosity of payments, and complementarity with child care policies, among other parameters, create in theory different incentives to have children and for women to work across countries, especially in the years immediately after childbirth. Consequently, for decades, countries have experimented with alternative PL policy designs.[1] Due to the documented persistent effects of motherhood on women's careers and the gender gap in earnings following childbirth (Bertrand et al. 2010, Angelov et al. 2016, Lundborg et al. 2017, Kleven et al. 2019a, Kleven et al. 2019b), how different designs of PL policies affect women's labor supply and fertility constitutes a critical question.

PL policies theoretically have the potential to have sizable effects on women's work decisions. Women with young children have the highest documented labor supply elasticities (Keane 2011, Attanasio et al. 2018), suggesting they are highly responsive to labor market incentives. Surprisingly, however, existing empirical studies almost all find negligible effects of even major PL reforms on women's labor market outcomes, especially three or more years after first birth.[2] Consequently, a growing number of economists has concluded that PL design has at best limited influence on women's decisions (e.g., Kleven et al. 2021).

In this paper, we offer new evidence on the effects of parental leave policy design on female labor supply and fertility. We make three main contributions. First, we show in a simple way that existing empirical studies – which mostly study women who gave birth around PL reform dates using a regression discontinuity design (RDD) – offer exceptionally clean identification, but estimate parameters that are difficult to interpret, and are potentially misleading to a policymaker. By comparing the ideal experiment to the typical quasi-experiment used in the literature, we document that by construction, the latter will underestimate – possibly severely – the effects of a PL reform. Additionally, we show that in the four European countries for which we have detailed survey and administrative data on PL take-up – U.K., Sweden, Austria, and Germany – women's different work and childbearing patterns are indeed strongly consistent with the specific incentives in their

[1]See Ruhm (1998), Lalive and Zweimüller (2009), Rossin-Slater et al. (2013), Schönberg and Ludsteck (2014), Carneiro et al. (2015), Dahl et al. (2016), Stearns (2017), Bailey et al. (2019), Farré and González (2019), Raute (2019), Kleven et al. (2021). For example, Dahl et al. (2016) consider a series of policy reforms in Norway that expanded paid leave duration between 1987 to 1992. Kleven et al. (2021), who study four major parental leave policy reforms in the 1990s and 2000s in Austria. Schönberg and Ludsteck (2014) evaluates five major expansions of maternity leave in Germany between 1979 and 1993. See Olivetti and Petrongolo (2017) for a review of family policies and its effects.

[2]See Baum (2003), Lalive and Zweimüller (2009), Schönberg and Ludsteck (2014), Carneiro et al. (2015), Dahl et al. (2016), Kleven et al. (2021), Corekcioglu et al. (2022).

countries' PL systems. These include both incentives that temporarily reduce labor supply in the year of or immediately after birth – e.g., by offering a highly paid parental leave – and incentives that increase labor supply shortly thereafter – e.g., by making parental leave compensation for second and third births a function of earnings prior to those births.

Based on this evidence, we then develop a structural model of women's labor supply and fertility decisions. Women make these decisions, as well decisions about consumption and savings, taking into account their effects on human capital accumulation and eligibility for parental leave benefits. The model's main features allow us to credibly validate the model and evaluate in detail the effect of different policy designs on women's decisions. First, as we will discuss in detail, the model allows us to simulate both the policy-maker's "ideal experiment," and the corresponding RDD quasi-experiment that is most commonly employed by researchers. As our second main contribution, we show that small RDD estimates of past policy reforms in the ten years after first birth, which our model replicates, are fully consistent with sizable true effects of the policy on employment and fertility. More broadly, a contribution of our approach is to illustrate the value of using a structural model to interpret a large body of exceptionally well-identified empirical estimates.

Finally, our model can simulate a range of detailed policy features, including duration of paid leave, duration of job protection, compensation structure, and eligibility rules. Additionally, our model incorporates other policy features that may be especially relevant to women's labor supply after birth, including child care policies and taxation structure. As our third main contribution, we evaluate in detail the effects of different PL design features, including those of the four countries in our data – U.K., Sweden, Austria, and Germany. In particular, we calibrate the model to U.K. data, and show that when we counterfactually expose U.K. women to Swedish, Austrian, or German policies, they make work and childbearing decisions similar to those of women in those countries. Notably, in these counterfactuals, preferences and social norms are held constant. The findings are consistent with women's significant responses to the strong incentives embedded in different PL systems.

Finally, we consider the set of model parameters that maximize women's labor force participation, fertility, and welfare, respectively. We show that a policy with a moderate duration and offering moderate to high wage replacement has positive effects on both fertility and labor supply.

The organization of the paper is as follows. Section 2 discusses existing empirical evaluations of PL reforms and why they estimate parameters that differ from those that a policy-maker needs to evaluate the reform's effect. Section 3 provides detailed empirical evidence on the PL policies as well as PL take-up, work, and fertility in U.K., Sweden, Austria, and Germany. Sections 4 and and 5 present the dynamic model and discusses calibration and model performance. Section 6 presents our main results and counterfactual policy experiments. Section 7 concludes.

2 Interpreting the Evidence on Parental Leave Policy

Economists have evaluated a number of major European parental policy reforms. Prominent examples include Schönberg and Ludsteck (2014), who evaluate five major expansions of maternity leave in Germany between 1979 and 1993; Dahl et al. (2016), who study a series of moderate and small policy reforms in Norway that expanded paid leave from 18 to 35 weeks between 1987 and 1992; and Kleven et al. (2021), who study four major parental leave policy reforms in 1990s and 2000s in Austria, which increased paid leave by up to a year. These and similar studies from other countries find that leave expansions somewhat reduce women's employment in the additional expansion period immediately following birth, and have negligible or no effects three or more years after birth (Baum 2003, Lalive and Zweimüller 2009, Schönberg and Ludsteck 2014, Carneiro et al. 2015, Dahl et al. 2016, Kleven et al. 2021, Corekcioglu et al. 2022). Dahl et al. (2016) and Kleven et al. (2021) also study subsequent fertility and similarly find negligible effects.[3]

The existing evidence on PL reforms is, by multiple measures, exceptionally high quality: it uses administrative data, major historical policy reforms, and credible research designs. In particular, because many PL reforms are implemented with strict eligibility cut-off dates, they present ideal settings for analysis using a regression discontinuity design (RDD). The typical empirical strategy can be illustrated with the following example: a PL reform extends leave duration from 1 to 2 years. It is announced in May 2000, and implemented on July 1, 2000. In this case, women who give birth in June are eligible for the old policy ("untreated"); those who give birth in July are eligible for an additional year of leave under the new policy ("treated"). Women who give birth in June and July 2000 are expected to be similar in most characteristics, other than their PL policy exposure. Moreover, since they were already pregnant when the reform was announced, there is little scope for women to choose to be on either side of the July 1 threshold.[4] Using panel data, researchers can follow the two groups of women before birth, to confirm that there are no differential pre-trends for the two groups. Then, they can follow them for many years after birth to estimate differences in their outcomes both in the short-run (one or two years after birth) and in the long-run (three, five, or ten years after birth). The estimates (see Figure 1) capture the causal effect of being eligible for an additional year of leave for the focal birth (in many studies, the first birth), compared the baseline leave amount.

While these estimated RDD treatment effects are well-identified, an important question is how

[3] Bailey et al. (2019) studies California's Paid Family Leave Act and finds negligible effects on completed child-bearing.

[4] The majority of studies also address possible remaining threats to identification with a range of robustness and sensitivity tests. Examples include leaving out a small window around the immediate cut-off date (July 1, in our example) to account for delayed inductions or elective C-sections; and addressing possible seasonal differences between June vs. July mothers by including additional non-reform years to control for seasonality.

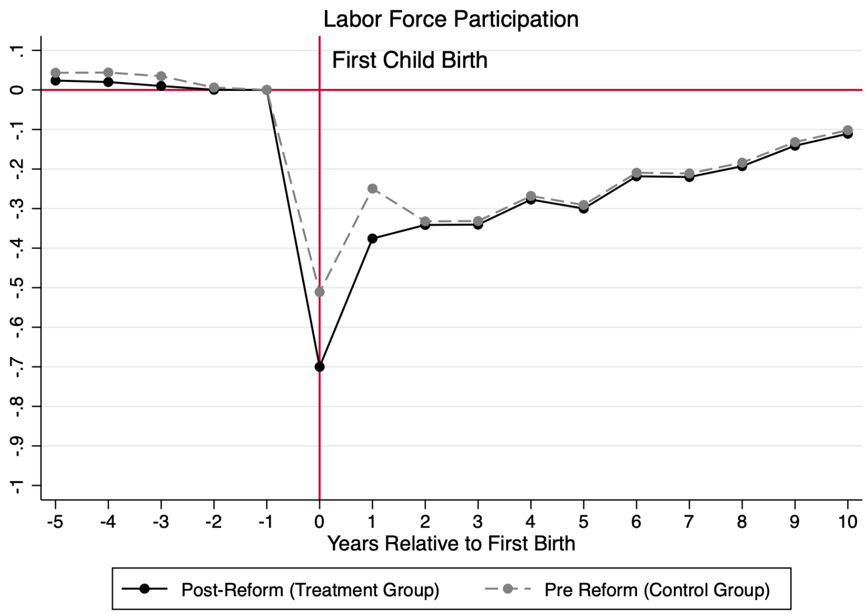


Figure 1: Stylized Example of Dynamic RDD Estimates

Notes: This stylized example illustrates the two most common features of dynamic RDD estimates of PL reforms: moderate effects immediately after birth (due to the change in duration in our hypothetical example), and virtually no long-run effects from around three years after first birth. For example, Kleven et al. (2021) find that changes in duration of the paid parental leave policy in Austria in the 1990s and 2000 generated around a 10% change in earnings due to its effect on labor supply in the first or second year after birth, depending on the reform. In all cases, they found no effects after year 3. Schönberg and Ludsteck (2014) find negative short-run effects on employment in Germany of between 10 and 30 percentage points, depending on the increases in the job protection and maternity benefit. However, they find that effects on employment three years after birth are at most around two percentage points, and typically close to zero.

they should be interpreted and, in particular, whether the parameters they estimate are informative to the policy maker about the main effects of the policy. To make progress on this question, we first consider the ideal experimental benchmark to evaluate a PL policy change. The ideal experiment, presented in Panel A of Table 1, is straightforward: it exposes two identical groups of women to different policies for all births (first, second, third, etc.), and also early in their lifecycle, since career decisions may be a function of expectations about job continuity and time out of the labor force after each birth. Comparing the two groups of women gives a policymaker information about the causal effect of the new policy on their outcomes, compared to the old policy.

By contrast, the typical RDD setting differs markedly from this ideal setting, as Panel B of Table 1 shows. Prior to birth, both groups of women are exposed to the old policy, and have identical expectations in early life. Thus, the policy's influence on labor market behaviors prior to first birth cannot be estimated. Second, both groups of women are under the new policy for potential second and third births. Their decisions about additional births and about work before

	Treatment Group	Control Group
Panel A: **Ideal experiment**		
Before Birth / Early Career	New	Old
First Birth	New	Old
Second birth	New	Old
Third birth, etc.	New	Old
Panel B: **RDD experiment**		
Before Birth / Early Career	Old	Old
First Birth	New*	Old
Second birth	New	New
Third birth, etc.	New	New

Notes: *New policy is a surprise when already pregnant with the child.

Table 1: Comparing New & Old Policy

and after those births are again governed by the same policy parameters and expectations. Finally, policy exposure differs for the first birth; however, a woman's treatment status is revealed only when she is pregnant or, if she is due close to the cut-off date, when she gives birth.

To summarize, given the limited scope for different behavioral responses by treatment status in the RDD quasi-experiment, small RDD estimates need not be inconsistent with large true effects of the policy. A reasonable interpretation of the evidence from RDD studies of PL reforms is that they offer exceptionally clean identification, but that the parameters they estimate are not, on their own, sufficient for policy evaluation. Consequently, the growing consensus that even large PL reforms have "virtually no impact" on women's labor market outcomes is premature. By construction, the RDD quasi-experiment shuts down all responses prior to birth; it is not well-equipped to capture different responses to the policy from three to ten years after first birth, when the majority of subsequent births occur, all under the same (new) policy environment; and, theoretically, it will capture well the effects on behaviors immediately after first birth only if women's decisions during this period are not affected by past decisions or by anticipated future births.

We conclude with two final points. First, the relatively few studies of PL reforms using classic differences-in-differences designs tend to find much larger effects of PL reforms on women's labor market outcomes. This is consistent with the idea that RDD estimates understate the true effects of the policy. For example, Ruhm (1998) finds that paid leave in nine European countries (Denmark, Finland, France, Germany, Greece, Ireland, Italy, Norway and Sweden) raised the female employment-to-population ratio approximately by 4%. Raute (2019) finds that a German maternity leave reform that raised the renumeration during leave especially for high-educated and high-earning women increased fertility of tertiary-educated women by 23%. A drawback of these studies is that it is more difficult to ascertain whether the selected control group provides a satisfactory counterfactual for trends in the treated group. For example, (Ruhm, 1998) uses men as a control group as well as a staggered cross-country design, which has important limitations. Raute (2019) uses low-educated mothers as a control group, since the reform primarily affected high-educated mothers. Despite less credible control groups, these studies more closely resemble the ideal experiment in that they maintain different policy exposures for the treated and control groups for all births, and also prior to birth. If the true long-run policy effects are sizable, these studies should find much larger estimates of these effects, compared to RDD studies, as they indeed do.

Finally, although the ideal experiment would yield the treatment effects that are most relevant to the policy maker, it is challenging to implement in practice. Similarly, it is difficult to conceive of quasi-experiments in which two otherwise identical groups of women are exposed to different policies over most of their lifecycles. Moreover, since PL policies are generally complex and can take many possible forms, evaluating the effect of each type of policy would require many treatment arms.

Therefore, in the remainder of the paper, we proceed as follows. First, to set ideas, we present descriptive evidence from four countries with PL policies that differ significantly in both length and compensation structure. We will discuss the incentives to work and have children in these policies, and show that women in those four countries make decisions about work, PL take-up, and fertility that appear consistent with incentives created by their countries' PL policy designs. Based on this evidence, we then develop a structural model of women's decisions, which can simulate a range of PL designs, and re-create both the ideal experiment to evaluate a PL policy and the RDD quasi-experiment. This approach will allow us both to study the effects of alternative PL policies and to evaluate credibility of our model, including its ability to replicate the small RDD estimates widely documented in the literature.

3 Evidence from UK, Sweden, Austria, and Germany

In this section, we present descriptive evidence from four countries for which have data on PL take-up: UK, Sweden, Austria, and Germany. Over the time period we consider, the PL policies differ significantly across these countries. After describing our data, we proceed in three steps. First, we discuss the main features of different PL policy designs, and discuss the policies in our four focal countries. Next, we discuss the incentives created by these policies, and evaluate whether the differences in women's labor supply and fertility are consistent with these incentives. Additionally, we discuss childcare subsidies and taxation, which may also influence the differences in outcomes we observe. Finally, we use our descriptive findings to discuss the types of policy responses that may be particularly difficult to evaluate using RDD quasi-experiments.

3.1 Data

Our data consists of household and individual panels from four European countries that all have, as a minimum, information on women's PL take-up, labor market activity, and timing of all births. For the U.K., we use the British Household Panel Survey (BHPS, 1991 to 2008). For Germany, we use the German Socioeconomic Panel (GSOEP, 1988-2015). For both Austria and Sweden, we use the Generations and Gender Survey (GGS). In Sweden, the GGS obtains information on individuals from administrative registers, covering the period 1992-2015. It also collects supplementary information via a household survey. In Austria, the GGS surveys individuals, who are asked to report each of their labor market activity statuses (employed, at school, on parental leave, homemaker, etc.) and their start and end dates since age 16, and covers the period 1991-2010.

Our main variables of interest in the descriptive analysis are the timing of the first and, if applicable, second and third births; and labor market status. We characterize the latter using three mutually exclusive and exhaustive categories: in the labor force, on parental leave, or home-maker. Rather than dropping women who are in school from our analysis, we designate them as participating in an economically productive activity and group them with women who are active in the labor force. In practice, however, this decision has little influence on the results.

In the data appendix, we provide details about how we construct the variables of interest in each dataset to ensure consistency. For analyses by time relative to first birth, we consider only women whose first birth occurred between 1993-2008. Additional restrictions, when necessary, are noted in the paper. For Austria, we consider years of first birth between 1996-2001 due to large policy changes before 1996 and after 2002. Similarly, for U.K., we consider years of first birth between 1995-2000 due to large policy changes after 2000. More detailed information about our data, sample, and variables can be found in Section A4. Table A1 shows the summary statistics for the women in our dataset.

3.2 Policy Design and Institutional Background

Historically, the key features of parental policy include the duration of paid leave; the duration of job protection; and the payment scheme. Payment schemes usually constitute either a relatively low fixed amount that is identical for all eligible individuals or, alternatively, a typically generous percentage of prior earnings. Consequently, we will focus on these main features of duration and pay in our analysis. Additional factors, often the target of more recent reforms, include whether some parental leave is earmarked for fathers, and whether a menu of flexible leave options is offered, such as an option for a shorter duration at higher pay. These are not the main objects of our analysis, although we will discuss them later in the paper.

Figure 2 summarizes the parental leave policies across our four countries of interest in our

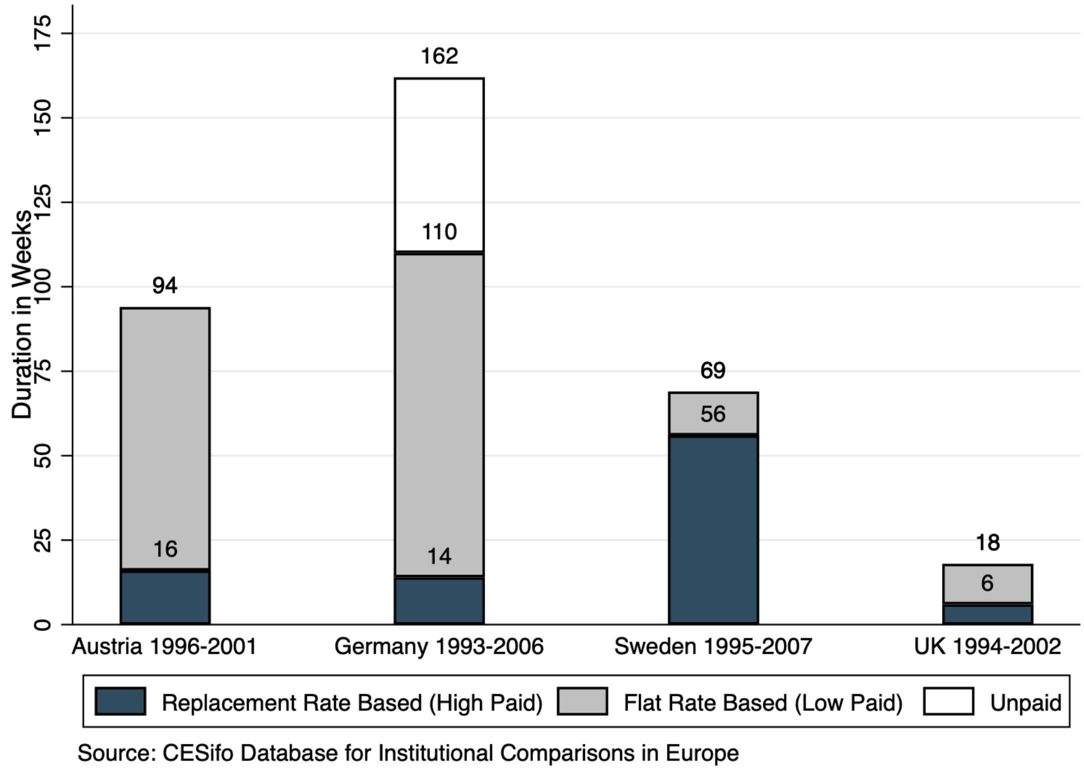


Figure 2: Parental Leave Policies in Europe

sample period. It provides a useful snapshot of common features and differences in PL policies across countries. For each country, we select a period of relative policy stability. We provide detailed histories of leave reforms and additional institutional details for each country in Appendix A3.

Figure 2 documents the maximum number of weeks of job-protected leave a mother can take in each country, and how many weeks are compensated at a high replacement rate, at a low fixed pay, or at no pay. In all these countries, most of the leave is government-paid. The generosity of fixed pay schemes across countries varies only mildly, at around \$500 per month. European countries typically offer around 14-16 weeks of high-paid maternity leave immediately around the time of birth, compensated at 100% wage replacement, following International Labor Organization (ILO) convention. The U.K. deviates from this convention somewhat, offering 6 weeks at 90% compensation. The remaining characteristics of leave, however, vary significantly across countries.

As Figure 2 shows, the U.K. has the least generous policy in our set of countries, with only 18 weeks of job protected leave, paid primarily at a low flat rate. Sweden has the most generous policy, at least as measured by compensation. It offers just over a full year of leave at 80% wage replacement, up to a high cap. Moreover, many establishments, through their collective agreements, privately "top up" this government-paid leave, so that some women receive nearly full pay during their leave. Germany and Austria offer the longest paid leaves, albeit at low pay, at around two years. On top of this, Germany offers a third year of job-protected, unpaid leave.

We have omitted some nuances in these policies, detailed in the appendix, but the above discussion provides a good summary of the most relevant features of these policies. Additionally, we note that eligibility for paid leave in all of these countries is contingent on prior employment. In Sweden, compensation is determined based on average earnings in the prior year, for all births. If a woman has a child within two years of the birth of the first child, she can take her second parental leave at the same compensation as the first one, even without working between the two births. Sweden in particular offers significant flexibility around when parental leave can be taken, as parents can take the leave up until their child's eighth birthday. While a fair share of fathers do indeed take their earmarked "daddy months" due to expire at the child's eighth birthday, Swedish mothers tend to take all or the majority of the leave immediately after birth (Ekberg et al., 2013).

Austria and Sweden both have earmarked time for fathers, which cannot be transferred to mothers. Therefore, this time is not included in the bar chart in Figure 2. Historically, countries have had mixed success with having fathers use the leave time allocated to them. For example, Ekberg et al. (2013) find that, following Swedish daddy month reforms, men's take-up of leave time increased, but was concentrated during the summer months and around the Christmas holiday. Bronson and Thoursie (2021) find more than half of men do not take their allocated "daddy months" within two years of the birth of their child, although a significant portion are likely to take it later (Ekberg et al., 2013). For these reasons, and because women continue to take the vast majority of the parental leave in the household, even when it can be transferred to the father, we will not focus on the leave taken by fathers in our analysis.[5] We refer readers to excellent work on paternity leave by numerous researchers, including Cools et al. (2015), Dahl et al. (2016), Bartel et al. (2018), Patnaik (2019), Farré and González (2019), Andresen and Nix (2024).[6]

3.3 Empirical Analysis

In this section, we document that not only the parental leave designs described above, but also transitions between work, PL, and home-maker status in the ten years after first birth differ significantly across the four countries in our sample. We also show that these differences are consistent with the incentives embedded in the respective PL systems.

Throughout the section, we will present our key outcomes of interest by time relative to first birth. To do this we adopt an event-study approach similar to that employed by Kleven et al. (2019a). We define the year an individual has her/his first child as $t = 0$ and label all other years

[5] See Appendix A2.1 for the results for fathers.

[6] See Goldin et al. (2020) for why competitive firms in the US provide paid PL both for fathers and mothers.

relative to that year. We follow individuals 5 years before and 10 years after first birth. Thus, t takes a value from -5 to 10. We use the following fairly standard specification, estimated separately for women and men in each country:

$$Y_{itk}^{gc} = \sum_{j \neq -1} \beta_t^{gc} \mathbb{1}[j = t] + \alpha_j^{gc} + \varepsilon_{it}^{gc}, \qquad (1)$$

where g corresponds to gender and c corresponds to country. In our results throughout this section, we plot the set of β_t^{gc} coefficients, which measure the change in the outcome variable relative to the year prior to birth, which is the omitted category. We control for age fixed effects, α_j.[7] The dependent variables are indicators for activity status (in labor force, on parental leave, or homemaker) or fertility status (has at least two children, has at least three children).

U.K. and Sweden

We begin by considering the U.K. and Sweden, which have arguably the most dramatic differences in PL designs. Recall that UK offers 18 weeks of job-protected leave at minimal pay, while Sweden offers slightly more than one year at 80% wage replacement. Before we examine the data, it useful discuss the incentives embedded in these policies.

The U.K.'s minimal policy is fairly straightforward and, because it is the least generous, effectively provides a useful baseline. Unless they have alternative private arrangements with their employer, U.K. women during our period of study generally must choose between returning to work approximately four months after birth or, alternatively, terminating their employment (becoming "homemakers" in our terminology) and searching for a new position once they are ready to return to the workforce. The PL system provides some incentives to return to work before having a second child, since eligibility for the next birth requires one year of employment prior to birth. However, as the policy offers a fairly short leave and low pay, these incentives may not suffice to outweigh the opportunity costs of time and, possibly, job search costs, if a woman terminated her employment after the conclusion of her first leave.

The set of incentives embedded in the Swedish system differ significantly from those in the U.K. Swedish women are effectively paid to stay home for just over a year, receiving nearly full wage compensation during their leave; this naturally incentivizes lower labor force participation compared to U.K. women during this period. Soon after the conclusion of leave, however, these incentives reverse. For one, Swedish women can exercise their option to return to their previous employer when their child is already one. Moreover to be eligible for a similarly highly paid parental leave for the second (or third) child, Swedish women need to be working prior to those

[7] Because we focus on fairly narrow time windows for first births, there is a high degree of correlation between time to birth, age, and year in our sample. For this reason, we do not additionally include year fixed effects.

higher-order births. Since the majority of mothers in Sweden – indeed, in all the countries we study – have a second child, they face significant incentives to return to employment soon after they conclude their first leave. For women planning to take another parental leave in the near future, the PL policy effectively acts as a major subsidy that nearly doubles her current wage. [8]

Figure 3: Women's Labor Market and Fertility Outcomes in Sweden and UK, by Years Relative to First Birth

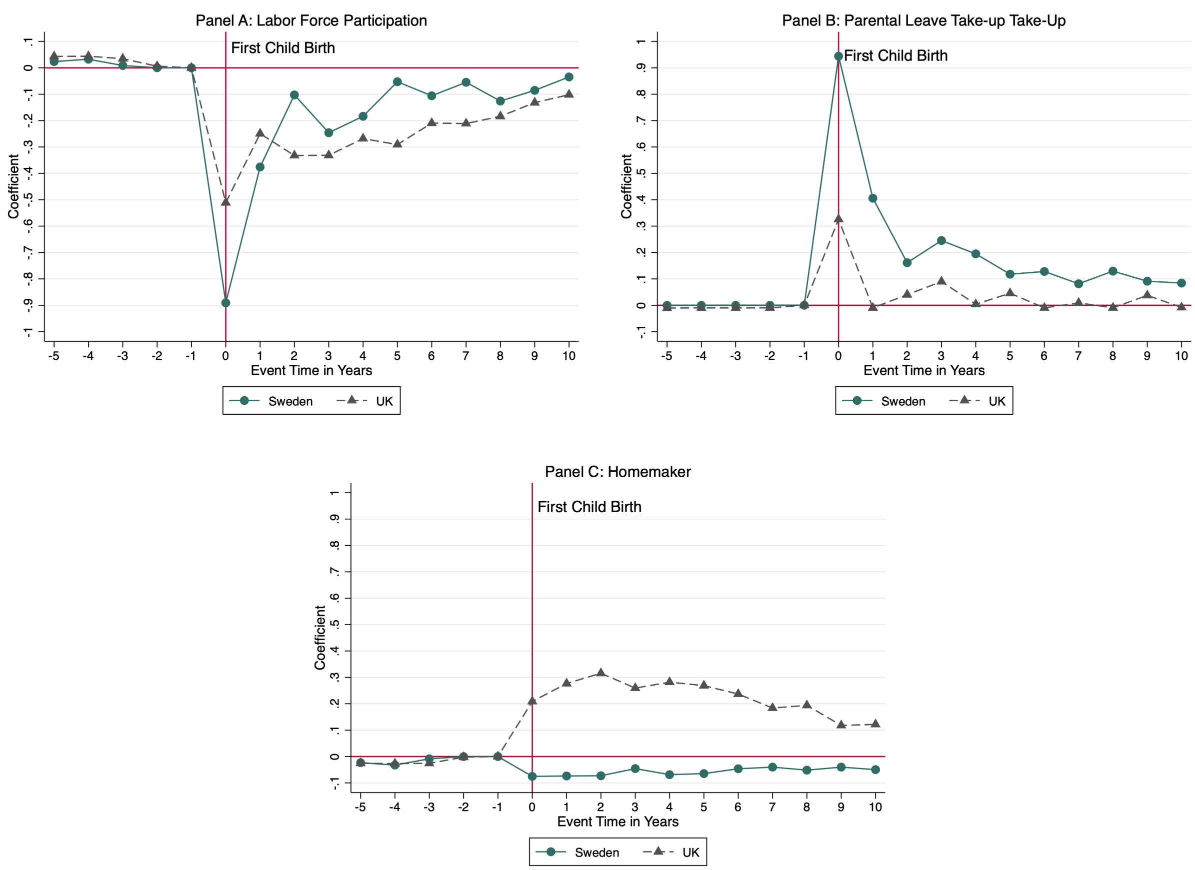
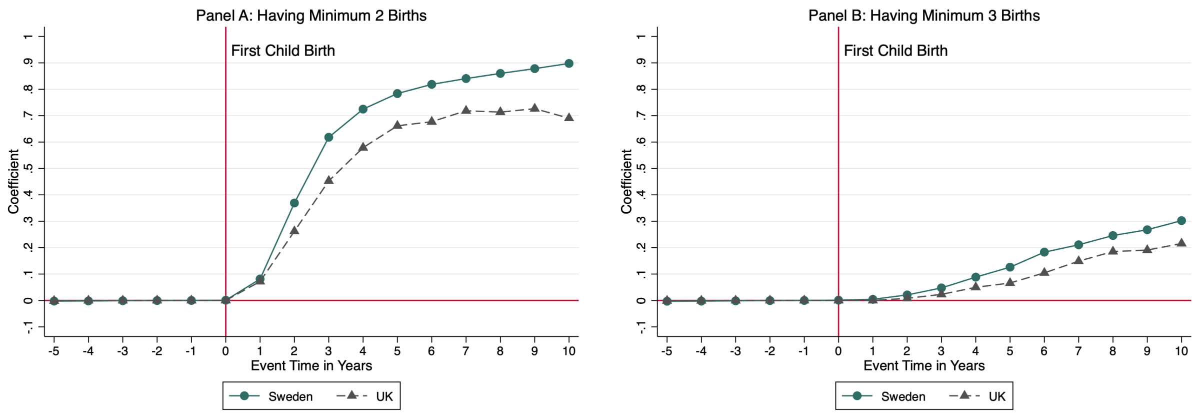


Figure 3 shows that differences in U.K. and Swedish women's work decisions are consistent with these different incentives. It features three main findings. First, panels A and B show that

[8]Since 1986, for consecutive births with less than 30 months spacing PL can be based on the parents' salaries prior to the birth of their previous child. Indeed, researchers who evaluated the 1986 reform found that the share of Swedish women with births in short succession increased following the reform, indicating that women responded to this incentive in PL system. However, it is important to note that this need not significantly diminish the incentives to return to work for women after an approximately one-year parental leave. First, to reduce the risk of health problems for both the mother and baby, it is medically recommended to wait 18 to 24 months after giving birth before becoming pregnant again. The official World Health Organization (WHO) recommendation is 24 months. Consequently, many women will elect not to take advantage of this feature of the PL system. Even for women who desire to have a rapid second pregnancy, there is medical uncertainty over whether she will be able to have a successful pregnancy and birth by the necessary cut-off date. Thus, unless a woman is already pregnant at the time she concludes her first parental leave, the incentive to return to work is nevertheless substantial.

Figure 4: Women's Fertility in Sweden and UK, by Years Relative to First Birth



virtually all Swedish women take parental leave immediately after first birth, consistent with the exceptionally generous PL compensation for those who take leave. [9] Consequently, the labor force participation of U.K. women exceeds that of Swedish women in the first 13 to 15 months following first birth. Though our data is not sufficiently fine-grained to evaluate monthly patterns, this is roughly reflected in the figure. Second, Panels A and B show that Swedish overwhelmingly return to work after the conclusion of their leave, consistent with the strong incentives to work prior having another birth. Indeed, for Swedish women Panels and B are virtual mirror images – over the ten years after first birth, Swedish women are either in the labor force or on parental leave. Lastly, the third main finding is that, by contrast, U.K. women transition at high rates from prior employment into a fairly persistent homemaker status (Panel C). More than 20% do so in the year they give birth, and this share increases further as women have higher-order births. Panel A shows that from about two years after first birth, Swedish women's labor force participation consistently exceeds that of British women. At the ten-year mark, the "motherhood penalty" in Sweden is still about 8 percentage points lower than in the U.K.

The patterns in Figure 3 are particularly striking since Swedish women also have more children U.K. women. Figure 4 shows that the shares of women who have at least two births and at least three births are significantly higher in Sweden. Nearly 90% of Swedish mothers have at least one more child within 10 years of the first birth, compared about 70% of U.K. mothers. Similarly, about 30% of Swedish mothers have a third child, compared to about 21% of U.K. mothers. This difference is also consistent with the PL designs. Swedish women face relatively low overall income loss due to first and higher-order births. Therefore, compared to U.K. women, they have more resources to have additional children and face fewer financial trade-offs per child.

[9] The levels of labor force participation rates in the year prior to birth (not pictured in the figure) slightly exceed 90% in both U.K. and Sweden.

Austria and Germany

Next, we add Austria and Germany to our comparison. Recall that both countries have long leave durations of approximately two years, paid at a low flat rate. Germany also offers women a third year of unpaid, but job-protected leave. Though the incentives relative to the U.K. and Sweden are slightly less clearcut in this case, a few general points can be made. First, on average, the policies incentivize longer leave-taking, by making it less costly to take an extended leave. Second, the financial incentives to return to work before having a second child are much weaker than in Sweden, since the policies offer a fairly low monthly pay of about $500. Third, the policy incentivizes fewer transitions to homemaker status, compared to the U.K., since women have a longer period to exercise the option to return to their previous employer. Alternatively, for those who have a second birth during the lengthy PL periods, the leave for the second child can be appended to the first leave.

Figure 5: Women's Labor Market Outcomes across Parental Leave Policy Designs, by Years Relative to First Birth

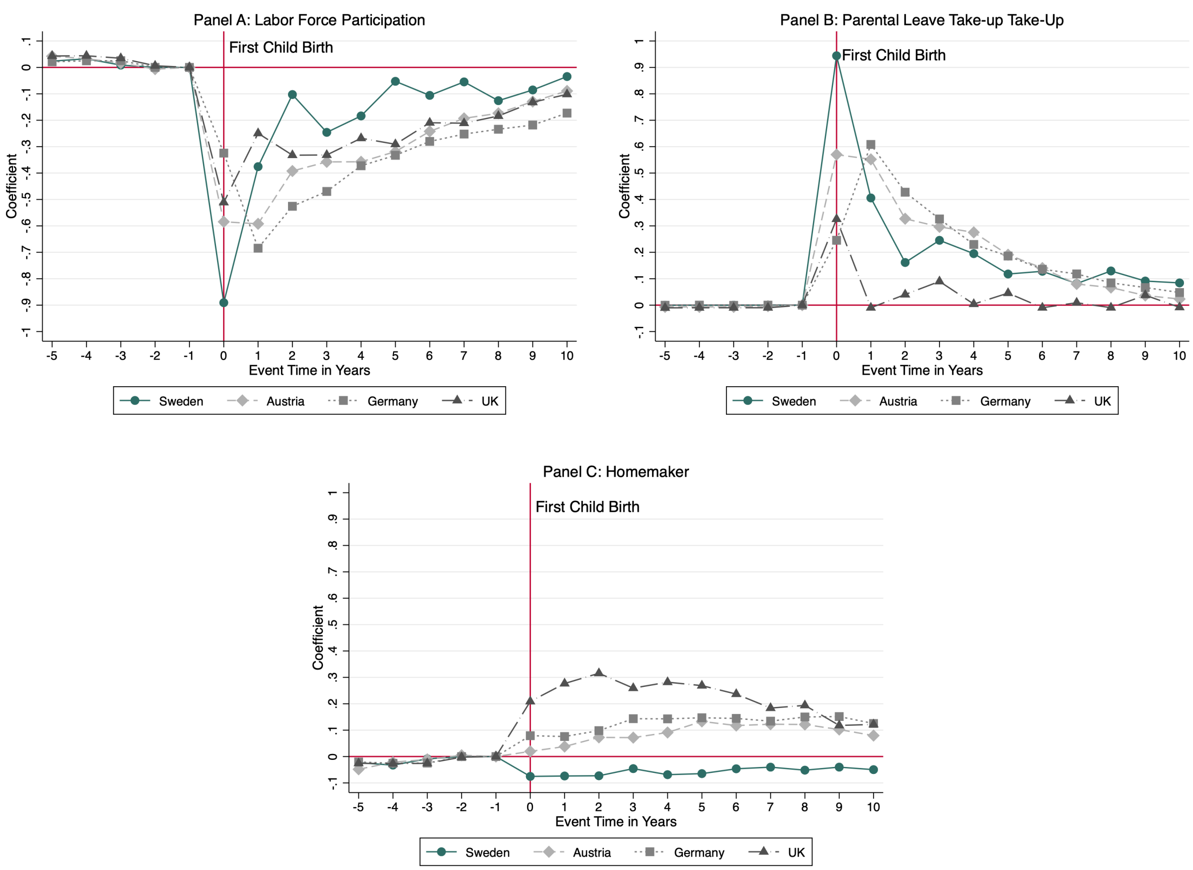


Figure 6: Women's Fertility Outcomes across Parental Leave Policy Designs, by Years Relative to First Birth

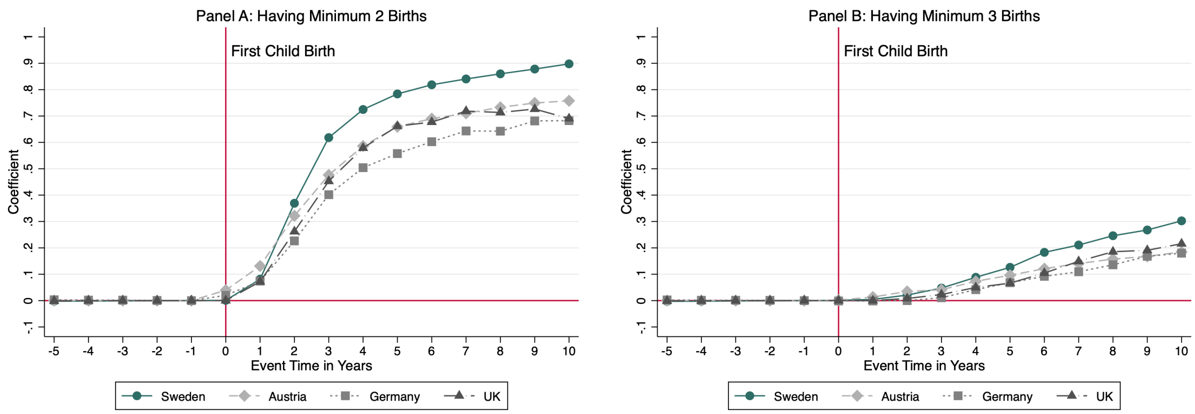


Figure 5 shows that women's decisions in Austria and Germany again are consistent with these incentives. The figure now compares all four countries. It shows that Austrian and German women do indeed take longer leaves, on average, around the time of first birth than women in the U.K. or Sweden. The additional year of unpaid leave in Germany appears to have relatively little impact. As expected, the rates of transition to homemaker status also fall between those in the UK and Sweden, as women either return to the labor force or commence a second lengthy leave. Finally, Figure 6 shows that rates of higher-order births are much lower than in Sweden and similar as in the U.K. Birth rates are even slightly lower in Germany than in the U.K. These fertility patterns are also broadly consistent with the features of the policy. Women work (and therefore earn) much less than in Sweden under these policies, and are compensated at only low levels during their leaves. Consequently, they have fewer resources and face greater financial trade-offs for each child, more akin to women in the U.K.

Taxation and Childcare

The countries we consider differ also along other dimensions that may affect women's decisions in the years following first birth. Two that are possibly important are the taxation system and child care policies.

Taxation. Differences in taxation – in particular, whether the tax system is individual or joint – have been found to have first-order effects on women's labor force participation (Bick and Fuchs-Schündeln 2018, Gayle and Shephard 2019). In particular, joint tax systems, which tax the pooled income of married couples, impose high marginal and average tax rates on secondary earners, and disincentivize secondary labor supply (Bronson and Mazzocco 2022). While this feature of the tax system is not affected by childbirth, differences in tax systems may have greater relevance for women with young children, since they are also the group with the highest labor supply elasticities.

While tax progressivity differs somewhat across the countries in our sample, three out of the four have individual systems: U.K., Sweden, and Austria. Germany has a joint system, which may help account for why mothers have lower labor force participation rates in Germany, despite having fewer second and third children. When we perform policy evaluation using our model, we will allow for different tax structures.

Childcare policies. Compared to the effects of taxation, the evidence on the effects of child care subsidies on women's labor force participation is more mixed. For example, Baker et al. (2008) and Lefebvre and Merrigan (2008) find sizable effects in Quebec, Bick (2016) finds small positive effects in Germany, while Havnes and Mogstad (2011) and Kleven et al. (2021) find that introducing childcare subsidies in Norway and Austria primarily crowds out informal arrangements, without affecting employment or other labor market outcomes.

Among the countries in our sample, only Sweden offers subsidized childcare for children ages one to two. When children are three to five years old, all countries in our sample offer free child care, primarily via pre-school. Appendix A3 provides further details. Since child care subsidies should increase take-home pay at least of women without grandparent care, they should theoretically influence labor force participation. Consequently, we will account for availability of subsidized child care in our model. Indeed, one possibly important design feature of PL systems is coordination with policies such as child care subsidies.

3.4 Summary

In this section, we documented several dimensions along which PL systems differ, and showed that they generate strong and highly heterogeneous incentives across the countries in our sample. Our event study analysis indicates that the different work and fertility decisions across the countries in our sample are also consistent with these different incentives. We conclude the section by noting which types of responses to PL policy incentives are particularly difficult capture using RDD quasi-experiments.

We have documented that by conditioning PL pay on prior work, many PL systems create incentives to work not just prior to first birth, but also potentially between births. In our sample countries, this was especially evident in the Swedish system, which uses a replacement rate payment scheme in which PL compensation depends on the prior year's earnings. Such anticipatory responses, however, are difficult for an RDD quasi-experiment to capture. To see this, consider a policy reform that moves a country from a UK-style policy to a Swedish-style policy. Suppose also that the true effect of the Swedish-style policy is to induce almost all women to work prior to the second birth. Since both the treated and control group in the quasi-experiment will be under the new Swedish-style policy for the second birth, they face the same strong incentives to be working again a year or so after birth, and differences between them are likely to be minimal. Focusing instead on second births near the reform date cut-off in the RDD quasi-experiment unfortunately does not in alleviate this problem – indeed, by design, it would again shut down all anticipatory effects prior to the second birth.

For closely related reasons, not just anticipatory responses – i.e., prior to each birth – but all medium- and long-run responses around higher-order births, are difficult for the RDD to capture. These subsequent births are exceptionally important in an analysis of motherhood penalties, as the arrival of (and responses to) second and third births largely determine observed outcomes three or more years after first birth. To show this explicitly, Figure 7 plots the share of women who are not in the labor force, by time relative to first birth and by the number of children. It shows that in each country in our sample, childhood penalties at least three years after the first birth are dominated by women having second and third births. Responses after all of these higher-order births are guided by the new policy for both the treatment and control groups in the RDD quasi-experiment. Returning to our previous example, this means that all women are now expected to take the year of almost fully compensated leave after their second or third births, regardless of whether their first birth occurred under a U.K.-style or Swedish-style policy.[10]

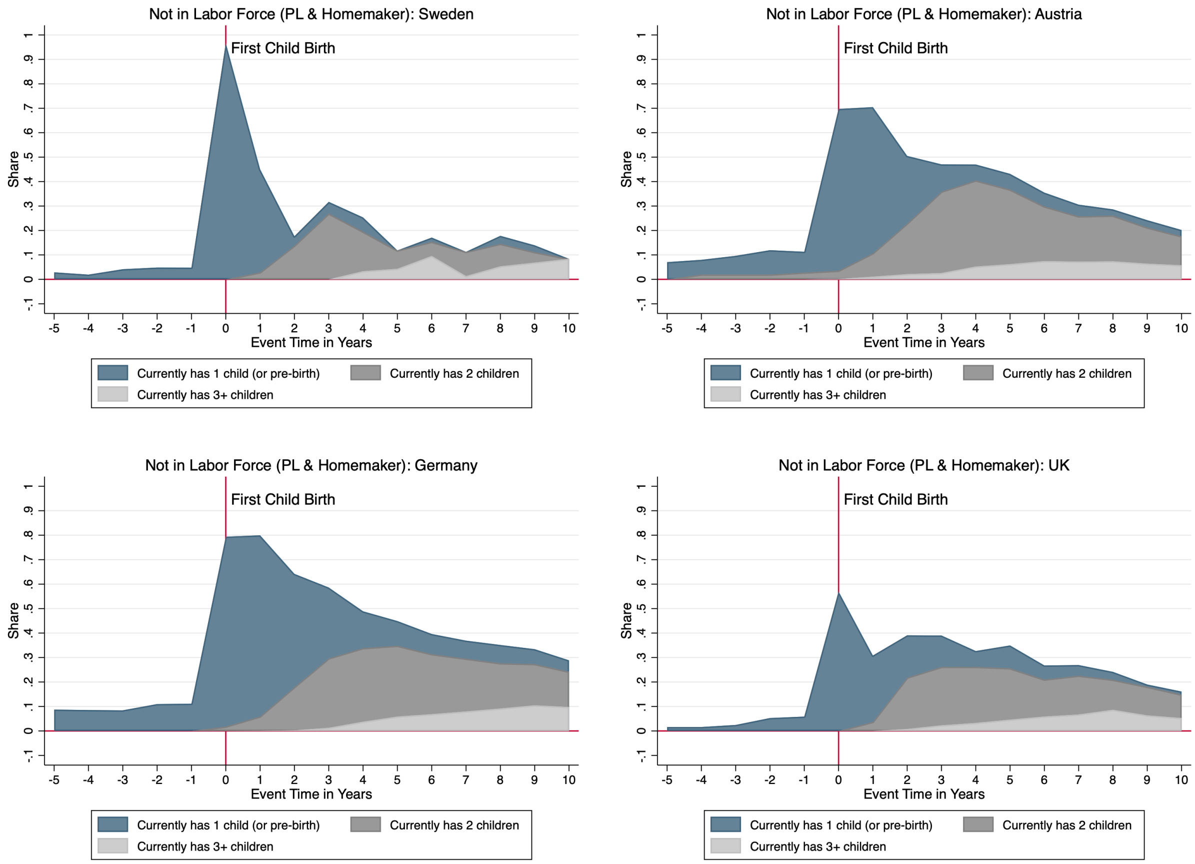
Our empirical analysis indicates that unlike the RDD quasi-experiment, the ideal experiment would fully capture the different responses to alternative PL policies, and yield the treatment effects that are most relevant to the policy maker. However, finding and/or implementing such an experiment in practice is challenging. Thus, in the next section, we develop a structural model of women's decisions, which can simulate a range of PL designs, and recreate both the ideal experiment to evaluate a PL policy and the RDD quasi-experiment.

4 A Dynamic Model of Labor Supply and Fertility

We build a dynamic model of individual decisions with the following main features. Women make decisions about labor supply (on both the intensive and extensive margins), about consumption and savings, and about whether to have children, and how many. Wages are endogenous and depend on human capital accumulation and depreciation. We model three important aspects of the institutional environment: the parental leave system, child care costs and subsidies, and taxes. The key elements of parental leave policy design that we incorporate into the model include the compensation structure, eligibility, job security provision and duration of leave. We include child care cost and subsidy amount and its availability by age. Lastly, we model the taxation system

[10] Moreover, they also face identical policy incentives to have those additional children. Thus, from year two or three years after first birth onward, the behaviors of the treatment and control groups are expected to look similar.

Figure 7: Share of Women Not in Labor Force, by Number of Children



(joint vs. individual) and the income tax schedule in each country that we study.

The model follows women from ages 22 to 50. In each period, they make decisions about labor supply h, time spent in home production $(1-h)$, consumption c, savings b, and whether or not to have a child. Productivity of time spent in home production depends on the number of children and age of the youngest child. Child care costs depend on age of the children and availability of (free) relative care. Time taken off affects wages via human capital accumulation and depreciation. Marriage, divorce, and the partner's income process are taken to be exogenous.

4.1 Preferences

Each period, individuals' utility is a function of consumption ($c_{i,t}$), hours worked ($h_{i,t}$), and number of children ($N_{i,t} \in \{0,1,2,3+\}$):

$$u_{i,t} = \frac{\left(\frac{c_{i,t}}{\psi_i(N_{i,t})}\right)^{1-\sigma_c}}{1-\sigma_c} + \beta_{i,t}\frac{(1-h_{i,t})^{1-\sigma_l}}{1-\sigma_l} + \kappa(N_{i,t}) + \phi(a_{i,t}, a_{i,t-1}). \tag{2}$$

The woman's consumption is expressed in per-adult equivalence units, where the adjustment term $\psi_i(N_{i,t})$ depends on marital status and number of children in the household. Utility from time spent not working in the labor market, $\beta_{i,t}$, depends on the number of children and the age of the youngest child, $a_{i,t} \in \{0,1,2,3,4,5,6+\}$. Households derive utility $\kappa(N_i)$ from having children, and disutility $\phi(a_{i,t}, a_{i,t-1})$ if a new birth occurs when an existing child in the household is zero or one year old.

4.2 Children

In each period, married women of childbearing age (under 40) choose whether or not they wish to have a(nother) child. Additionally, both married and unmarried women can have an unplanned birth. The probability of both planned and unplanned births, $q_p(age)$ and $q_u(age)$ respectively, declines with the age of the mother, and is zero after age 40.

Children affect household utility via three main mechanisms. First they require resources: $\psi_i(N_{i,t})$ increases with the number of children, and captures that households with more children require more resources to obtain the same level of per-adult-equivalent consumption. Additionally, households with young children pay a child care cost, $K(a_{i,t}, G_i) \cdot h_{i,t}$, which depends on the mother's hours worked, the age of the youngest child, the institutional environment (e.g., free or subsidized child care, by age), and whether the household has access to free grandparent or other relative care ($G_i \in \{0,1\}$). Fraction ξ of households have access to such free care. This latter feature is necessary to capture that a large share of households with young children pay no child care expenses even when children are too young to be eligible for free childcare in a given country, and both adults work.

Second, children affect the woman's utility from time spent outside of labor market work by shifting $\beta_{i,t}$, which depends both on the number of children and age of the youngest child:

$$\beta_{i,t} = \gamma_0 \qquad \text{if } N_{i,t} = 0 \tag{3}$$

$$\beta_{i,t} = \gamma_1 \qquad \text{if } N_{i,t} > 0, a_{i,t} = 0 \tag{4}$$

$$\beta_{i,t} = \gamma_2 + \gamma_3 \cdot (5 - a_{i,t}) + \gamma_4 \cdot N_{it} \qquad \text{if } N_{i,t} > 0, 0 < a_{i,t} < 6 \tag{5}$$

$$\beta_{i,t} = \gamma_5 + \gamma_6 \cdot N_{it} \qquad \text{if } N_{i,t} > 0, a_{i,t} \geq 6 \tag{6}$$

For $\gamma_k > 0$, this parameterization captures that utility from home production increases with the number of children, and is greater when children are younger. The separate parameter for the year of birth ($a_{i,t} = 0$) provides flexibility to capture that disutility from work is likely to be exceptionally high in that year, compared to years when the youngest child is already at least one year old.

Next, children generate direct utility for the household through the term $\kappa(N_i)$:

$$\kappa(N_t) = \sum_{j=0}^{3} \kappa_j \mathbb{1}(N_{i,t} = j)$$

where κ_0 is normalized to zero for households without children.

Lastly, households experience the utility cost $\phi(a_{i,t}, a_{i,t-1}) = \phi_0$ only if they had a child age 0 or 1 in the previous period ($a_{i,t-1} \leq 1$), and have a child again in the current period ($a_{i,t} = 0$). Without this parameter, the model would predict counterfactually small spacing between births. It captures in a simple way a multitude of factors that influence the desire not to have children in overly rapid succession: desired or required recovery period for the mother; preferences to continue breastfeeding the current child; reduced fertility during the breastfeeding period; and time demands in caring for two children under age two, among other possible reasons not easily captured elsewhere in the model.

4.3 Wages, Spousal Income, and Budget Constraint

Women's wages depend on human capital accumulation:

$$w_{i,t} = w_{i,t-1} \cdot (1 + \alpha h_{i,t-1}) + \varepsilon_{i,t} \text{ if } h_{i,t-1} > 0 \quad (7)$$

$$w_{i,t} = w_{i,t-1} \cdot (1 - \delta_t) + \varepsilon_{i,t} \qquad \text{if } h_{i,t-1} = 0 \quad (8)$$

where $\varepsilon_{i,t} \sim N(0, \sigma_f^2)$. Leaving your job ($h_{i,t} = 0$) generates a wage penalty of δ_{t+1} in the following period.

Individuals marry a spouse each period with a probability $p_m(age_i)$ that depends only on the age of the woman. Similarly, among married couples, divorce occurs with probability $p_d(age_i, N_i)$ each period, and depends on age and number of children. If the woman married, her spouse's income, $y_{i,t}$, evolves exogenously with time:

$$y_{i,t} = y_{i,t-1} + \eta_{i,t} \quad (9)$$

where $\eta_{i,t} \sim N(\mu_t, \sigma_m^2)$, where μ_t is average wage growth in period t.

The budget constraint faced by the household is characterized by:

$$\tau(w_{i,t}h_{i,t}, y_{i,t}) + PL_{i,t} + s_{i,t} = c_{i,t} - K(a_{i,t}, G_i)h_{i,t} + s_{i,t+1} \quad (10)$$

where $\tau(w_{i,t}h_{i,t}, y_{i,t})$ is the household's after-tax income and $s_{i,t}$ is savings. $PL_{i,t}$ is the parental leave income.

4.4 Parental Leave

Parental leave income, $PL_{i,t}$, is positive if an individual has a child that period and decides to go on parental leave ($h_{i,t} = 0$), zero otherwise. PL compensation depends on the policy adopted:

$$PL_{i,t} = 0 \quad \text{under an unpaid system} \tag{11}$$

$$PL_{i,t} = \overline{PL} \quad \text{under a fixed/flat rate system} \tag{12}$$

$$PL_{i,t} = w_{i,t-1} \cdot h_{i,t-1} \cdot r \quad \text{under a replacement rate system} \tag{13}$$

where r is the rate of replacement. Other aspects of the policy can be easily embedded in the framework. Eligibility criteria are modeled by restricting PL to those who worked in the prior period, $h_{i,t-1} > 0$. Job security is incorporated by setting wage penalty $\delta_t = 0$ during parental leave. This captures the idea that women can return to work at their previous wage after the conclusion of parental leave, but do not accumulate any experience.

4.5 Women's Decision Problem

We can now describe decision problem of a woman. The woman enters a given period with state variables $\omega_t = (age_{i,t}, w_{i,t-1}, h_{i,t-1}, y_{i,t-1}, s_{i,t-1}, N_{i,t-1}, a_{i,t-1}, G_i, m_{i,t-1})$, corresponding to her age and the previous period's wage, hours worked, savings, number of children, age of the youngest child (if children are present), whether she has access to free grandparent care ($G_i \in \{0,1\}$), and whether she was married in the previous period ($m_{i,t} \in \{0,1\}$).

She chooses consumption $c_{i,t}$, hours of work $h_{i,t}$, savings $s_{i,t}$ and ideal number of children N^* in the following period that maximize her life-time expected utility.

$$V_{i,t} = max_{c_{i,t},h_{i,t},s_{i,t},N^*_{i,t}} u_{i,t} + \beta E[V_{i,t+1}(\omega_{t+1}|\omega_t)] \tag{14}$$

$$\text{st. } \tau(w_{i,t}h_{i,t}, y_{i,t}) + PL_{i,t} + s_{i,t} = c_{i,t} - K(a_{i,t}, G_i)h_{i,t} + s_{i,t+1} \tag{15}$$

$$0 \le h_{i,t} \le 1 \tag{16}$$

$$N_{i,t} = g(N^*_{i,t}, N_{i,t-1}, age_i, m_i) \tag{17}$$

$$w_{i,t} = f(w_{i,t-1}, h_{i,t-1}) \tag{18}$$

$$y_{i,t} = y_{i,t-1} + \eta_{i,t} \qquad \text{if } m_i = 1 \tag{19}$$

$$y_{i,t} = 0 \qquad \text{if } m_i = 0 \tag{20}$$

The constraints include the budget constraint, time constraint, the fertility process that determines the actual number of children, conditional on desired and existing number, the process for the evolution of her wage and, if married, her spouse's income. $E[V_{i,t+1}(\omega_{t+1}|\omega_t)$ is the expected value function of the woman when she enters period $t+1$.

5 Model Calibration and Performance

We calibrate the model using British Household Panel Survey (BHPS) data. We target three key sets of moments related to fertility and labor supply. The moments related to fertility include: the share with 0, 1, 2, and 3+ children ever born and the age at first birth (5 moments); and the probabilities of second and third birth, 1, 2, 3, 5, and 10 years after first birth (10 moments). The second set of moments targets women's labor supply: employment and hours worked by age of youngest child and total number of children (22 moments). The third set of moments is a miscellaneous set required to obtain parameters related to the share with grandparent care, and the parameters of the wage and spousal income processes. It includes the share that work and rely on care by relatives; and men's and women's wages, by work history. We calibrate women's fecundity by age to data from the medical literature (Khatamee and Rosenthal (2002)). We calibrate unplanned births for single individuals and couples to those observed in BHPS. We assign the utility function parameters $\sigma_c = 2.1$ and $\sigma_l = 1.9$. The whole set of parameters used for calibration is shown in Table 2.

Parameter description:	Paramater	Value
CRRA paramter	σ_c	2.1
Power parameter on leisure/home production	σ_l	1.9
Coefficient on disutility from work (no children)	γ_0	0.68
Coefficient on disutility from work (child age 0)	γ_1	1.11
Coefficient on disutility from work, constant term I	γ_2	0.76
Coefficient on disutility from work, effect of child's age	γ_3	0.02
Coefficient on disutility from work, effect of no. of children I	γ_4	0.01
Coefficient on disutility from work, constant term II	γ_5	0.66
Coefficient on disutility from work, effect of no. of children II	γ_6	0.03
Utility from having one child	κ_1	0.82
Utility from having two children	κ_2	0.93
Utility from having three children	κ_3	1.02
Disutility from closely spaced births	ϕ	-1.03

Table 2: Utility Function Parameters for Calibration

Next, we show the model performance/fit by comparing the model results with the data from British Household Panel Survey. Figure 8 graphs women's labor force participation and her probability of having a second and third birth, relative to the time of first birth. The model parameters imply very high utility from time spent not working in the year after first birth, which generates the large and well-known drop in women's labor supply following first birth in Figure 8, even though UK women have only very short parental leave in the time period we consider, in the 1990s. The model then successfully replicates a small uptick in the following year, as some women return to work, with the remaining share of women temporarily leaving the labor force, given that utility from time spent at home continues to be high while children are young. The share working again dips in years 2 to 5, when virtually all second births occur, as some of the women who did return to work after the first birth again reduce their labor supply for the second birth. Return to work is a relatively slow process, partly because some women continue on to have a third child, and partly because wages for women deteriorated during their absence(s) from the labor force, reducing incentives to work in the long-run. Even ten years after birth, only about 80% of those who were originally working in the years prior to birth return to work.

The model parameters related to utility from children, κ_1, κ_2, and κ_3, imply that the largest jump in utility is between having no children and having one child. Even though the additional utilities to having a second child ($\kappa_2 - \kappa_1$) and a third child ($\kappa_3 - \kappa_2$) are roughly similar, only a small share of women has three children, since the penalties incurred for labor force absences amplify significantly over time, with each additional child. Moreover, children require resources,and the household becomes more resource-strapped as women spend more time out of the labor force, reducing the desirability of a third child.

6 Main Results

In this section, we present the main set of counterfactuals we will use to study the effects of different parental leave policy designs on women's labor supply and fertility decisions. First, we simulate and evaluate the different institutional designs of the three other countries we have discussed at length in this paper: Sweden, Austria, and Germany. Second, we compare the policy-maker's ideal experiment and the RDD quasi-experiment for actual historical policy reforms, since we can simulate both benchmarks using our model. Third, we isolate the effects of individual components of PL policies, as well as their interactions and relationship to childcare policy, on long-run labor supply. Finally, we conclude by identifying the sets of parameters that maximize labor force participation, fertility, and welfare in our model, and discuss the trade-offs a social planner faces when maximizing each of these variables.

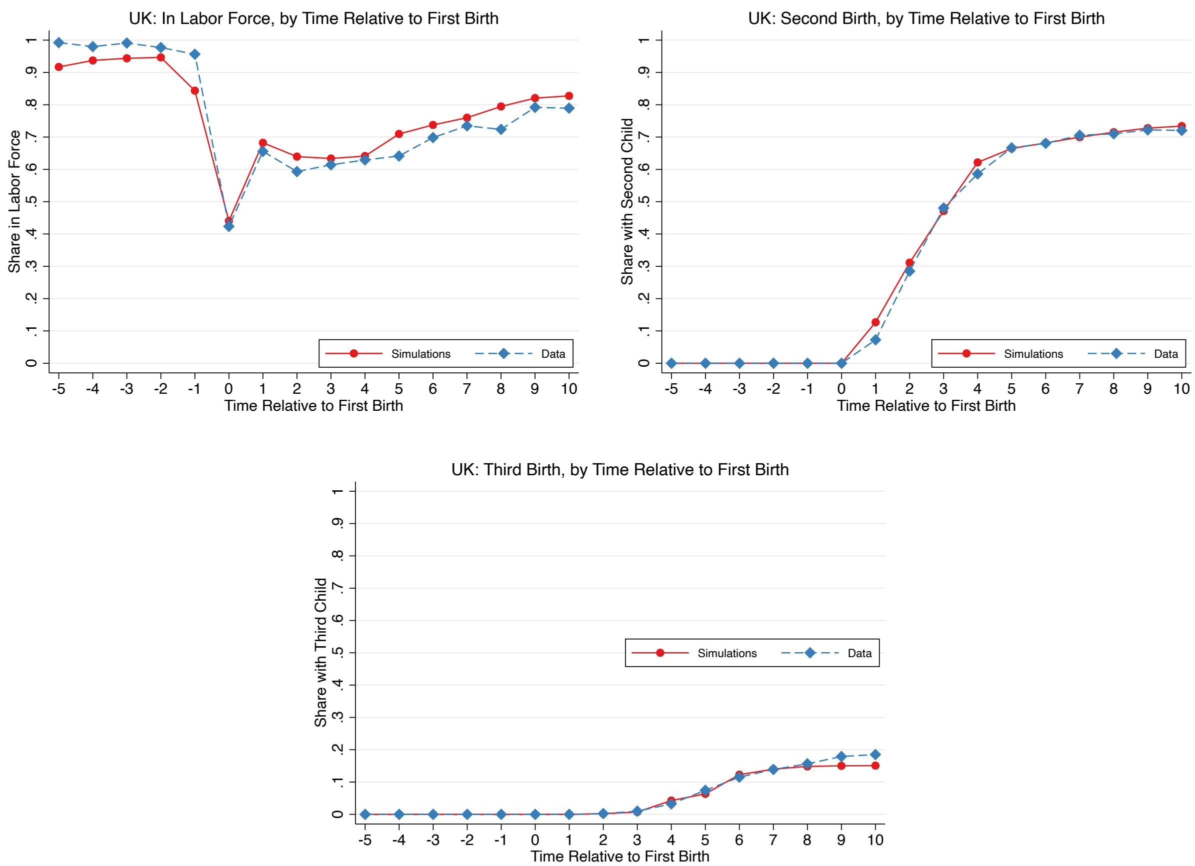


Figure 8: Women's Labor Force Participation and Having a Second and Third Birth Relative to First Birth: Model Results and British Household Panel Data

6.1 Baseline Counterfactuals: Simulating Sweden, Austria and Germany

In our first set of counterfactuals, we simulate the different institutional designs of the three other countries we have discussed in detail in this paper: Sweden, Austria, and Germany. Specifically, we simulate the parental leave policy designs in those countries – the duration of paid leave and job protection, compensation, and eligibility rules – along with childcare and taxation policies. Since our model is calibrated to UK data, these counterfactuals effectively answer the following question: how would U.K. women's decisions change if they were assigned the institutions of other countries? Naturally, we do not expect U.K. women's responses to exactly match the data from the respective countries: preferences, norms, and other policies not captured by our model may all differ across countries. However, we will show that the simulations align with the actual decisions of women in Sweden, Austria, and Germany remarkably well.

Sweden. We begin by simulating Sweden's institutions in our model. We present the results in Figure 9. Each panel plots three series, corresponding the U.K. baseline (in solid red); the

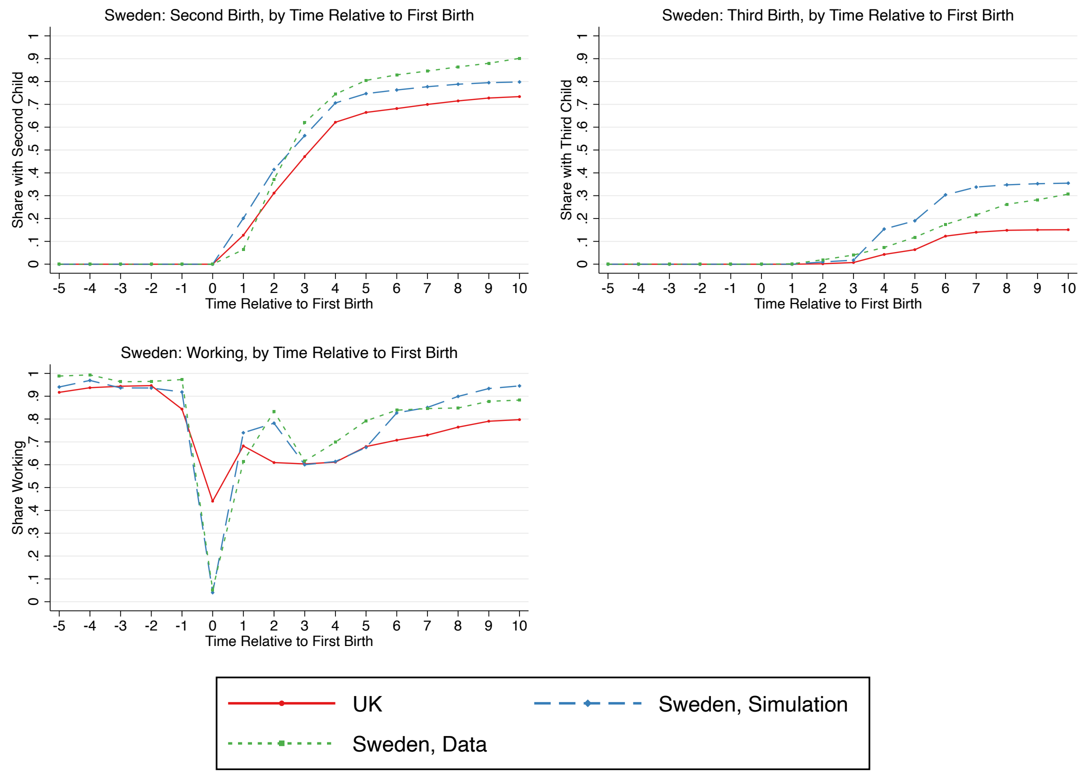


Figure 9: Baseline Counterfactual: Sweden

simulation of U.K. women under Swedish institutions (dashed blue); and the Swedish data (dotted green). To simulate the parental leave policy, we follow the coding described in Section 4.4. Recall that most of the leave in Sweden is based on a replacement rate scheme where mothers receive 80% of their pre-birth earnings during their leave. Our simulation predicts that under this policy, virtually all women take up the parental leave in the period they give birth, and almost immediately return to work upon its conclusion. This is consistent with the strong incentives in the PL design: in the year they give birth, women still receive 80% of their pay if they do not work; in the year after birth, every dollar they earn is effectively worth 1.8 dollars, since most decide to have a second child very soon after the first. The fact that the model produces responses that are so similar to the data is impressive, but should not be altogether surprising. In a standard dynamic labor supply model such as ours (e.g., Keane (2011)), these policy-induced taxes and subsidies to women's wages are of quite exceptional magnitude, and expected to have sizable effects.[11]

When we remove Sweden's childcare subsidies for children ages one and two, we find that average employment in the first three years after birth decreases by less than 1.5 percentage points. Even though subsidized childcare acts as a further wage subsidy, its effect on employment is small,

[11]Recall that ours is by design quite simple, and does not deviate significantly from the classic model.

since the PL policy already generates sufficient incentive to return to work when women's youngest child is one or two years old.

Lastly, we find that the model predicts higher rates of second and third births under the Swedish policy, as Figure 9 shows. Under the Swedish policy, U.K. women lose significantly less income over the 10 years after first birth – they are fully compensated during their leave, and also more likely to continue working. This means that U.K. women have more resources to have additional children, and also face fewer financial trade-offs for each additional child.

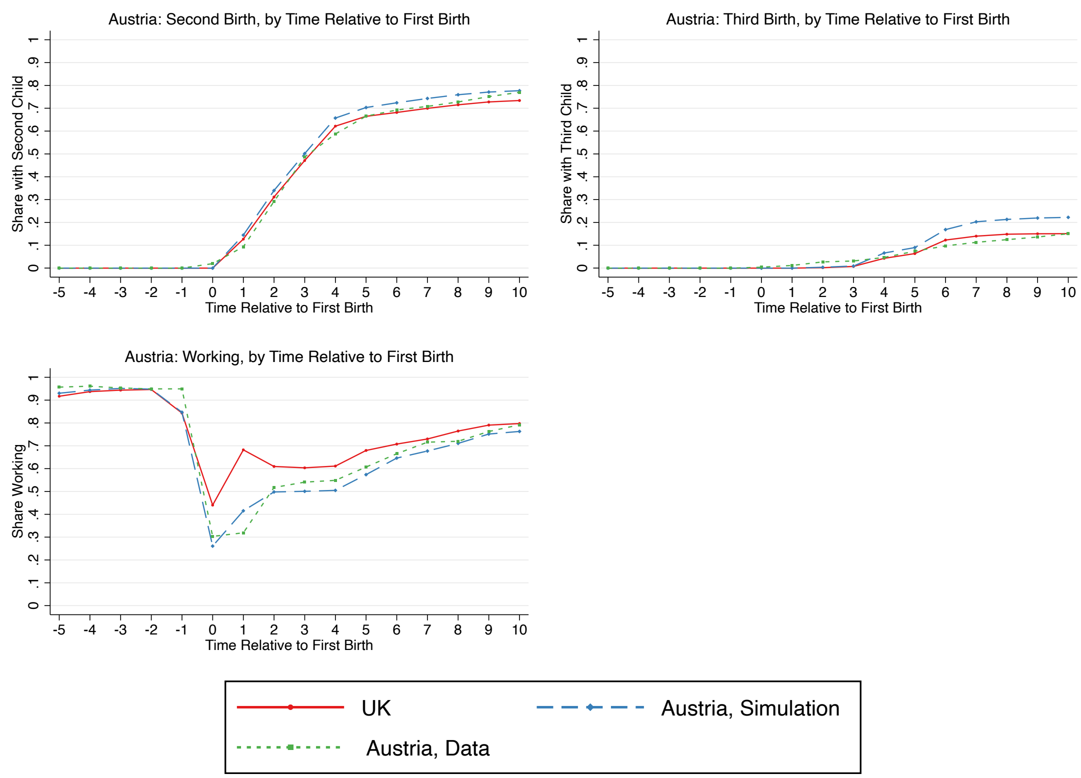


Figure 10: Baseline Counterfactual: Austria

Austria. Next, we consider Austria in Figure 10. For Austria, we model the policy as a two-year policy which pays everyone who was working in the year prior to birth $500 per month. In line with the policy, we allow women who work part-time to receive partial benefits. Our model again predicts that under Austrian institutions, U.K. women behave quite similarly to Austrian women. First, it predicts substantially longer leave-taking. Given the low PL compensation, not all women take even a full year of leave. At the same time, more than half take two years. Our model parameters indicate that women value the time off with children under two years old especially highly. The pay and in particular the ability to secure the same wage after even two years off induces more women to take a lengthy leave.

The model predicts that effect of the policy on higher-order births for U.K. women is small. In principle, the Austrian policy is strictly more generous than the U.K. policy, which should increase fertility. However, the Austrian policy does not generally increase the resources available to households, since more time is spent out of the labor force and PL compensation is low. Consequently, the predicted fertility response is much smaller than under the Swedish policy.

Germany. We report the simulation results for Germany in Section A5 for brevity. The results are similar to Austria due to substantial overlap in the institutional designs.

To conclude, our model predicts outcomes for U.K. women under the policy designs of other countries that are very similar to those of their counterparts in those countries. Theoretically, they need not be the same: different preferences, social norms, and other policies not captured by our model could all generate a significant gap between U.K. women's predicted behaviors under, say, Swedish institutions and actual observed outcomes in Sweden. However, the fact that the model predictions for U.K. women are so close to the data from other countries is consistent with the substantial incentives generated by PL designs, during a period when women's labor supply elasticities are also highest. Additionally, it indicates that social norms – – while potentially important for passing these policies – are on their own not necessary to account for the observed differences in patterns across countries (e.g., Kleven et al. 2021).

6.2 Policy Reform Counterfactuals: Simulating the RDD Setting

Next, we consider a counterfactual that allows us to compare results generated by our model to an RDD-style quasi-experiment. Our first example focuses on the Austrian reform in 2002 that increased PL from two to three years and removed all work eligibility requirements for PL pay. Our second example focuses on a hypothetical policy change that moves a country from the U.K.'s institutions to Sweden's.

We begin with the 2002 Austrian reform. Simulating the ideal experiment is straightforward: it simply requires simulating our main model under the old policy; simulating it again under the new policy; and comparing the two series. Simulating the RDD quasi-experiment in the model requires an additional step. As before, we simulate the model under the old and new policies, saving the value function approximations for each. Women in the model make decisions with expectations based on the pre-reform environment up until the year of birth; thereafter, they make decisions with expectations based on the new post-reform environment. The treatment and control groups differ in their policy parameters for the first period only.

Our results are summarized in Figure 11. The first graph compares the two policies under the ideal experiment. The two series differ noticeably, with a significant drop in labor supply post-

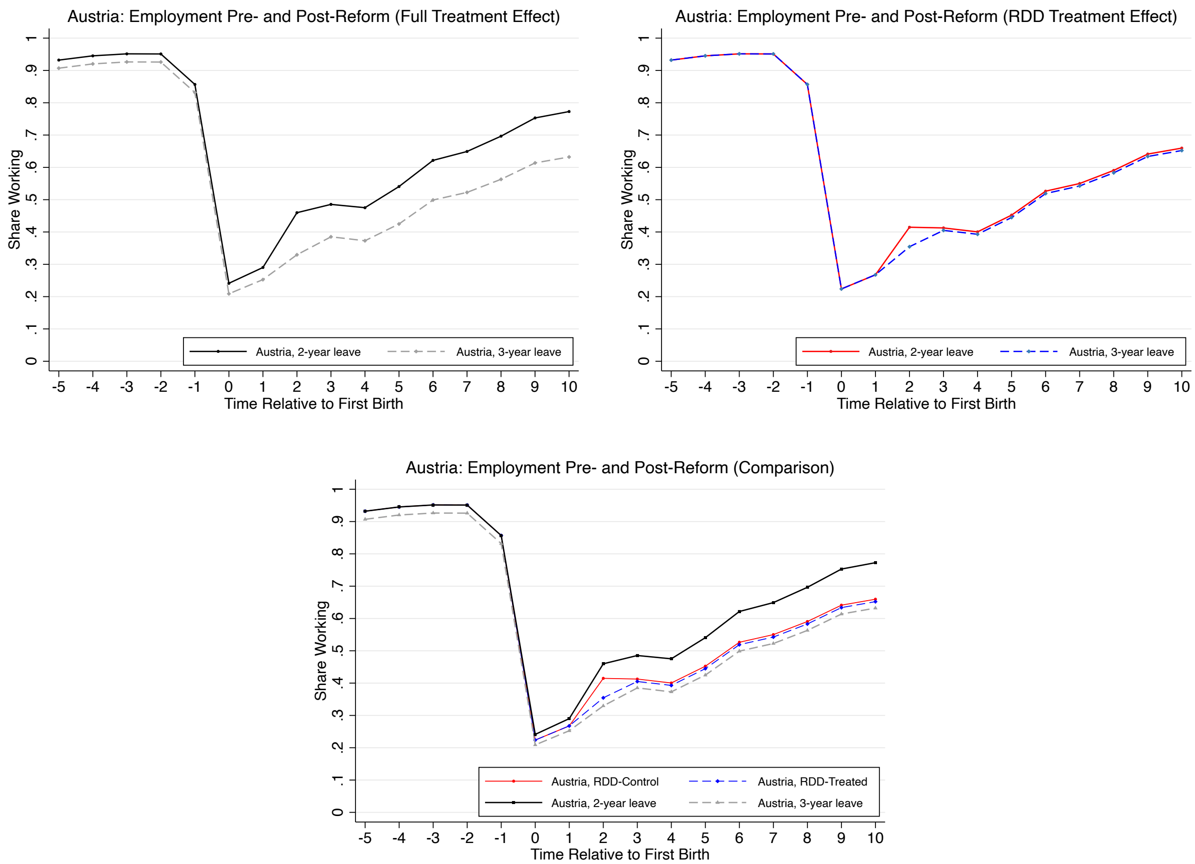


Figure 11: Policy Reform Counterfactual and Simulating the RDD Effect

reform. Since the duration of job-protected leave is extended by another year, more women to take it up. Additionally, eligibility requirements based on prior employment are dropped entirely, reducing work incentives both prior to birth and, more relevantly, between births. The effect prior to first birth is small, since most women have low opportunity costs during this period and therefore work at high rates.

The remaining two graphs in Figure 11 visualize, respectively, the simulation results for the effect of each policy under the typical RDD quasi-experiment; and all four series together. We highlight three main results. First, the RDD "treated" group – which is exposed to the new policy starting from the first birth – captures quite well women's decisions in the post-reform environment. This is not surprising, since all births for this group are under the new policy in both the ideal experiment and the quasi-experiment. While the RDD (by design) cannot pick up the effect prior to first birth, this effect is overall small. The second main finding is that the RDD "control" group very poorly captures the decisions of women under the old policy. This is also not surprising, since higher-order births for these women are under the new policy; moreover, RDD-control women's expectations about the future update immediately, including about the the removal of work requirements for future leaves. Consequently, even one and two years after first birth, the outcomes of women in RDD control group differ from those who are fully under the two-year policy. As a result, the RDD quasi-experiment in the model generates only small effects in the short-run, with negligible long-run effects, the third main result.

6.3 Parental Leave Parameters that Maximize Female Labor Force Participation and Welfare

In our final set of policy counterfactuals, we evaluate which PL parameters maximize, respectively, welfare and female labor force participation.[12] We ensure that policies are revenue-neutral by adjusting marginal tax rates by an identical increment in all brackets. Additionally, we assume free universal childcare starting at age 3, consistent with the policies in most European countries, and individual taxation. The parameters we consider include duration of paid leave and duration of job protection. Each of these can take values of 0, 1, 2, or 3 years. We use increments of one year because this is the length of a period in our model. We allow for a function for the payment scheme that nests pure wage replacement, fixed pay, or a possible combination of both. We consider replacement rates in 5% increments, and fixed pay amounts in $500 increments. We consider three types of eligibility criteria: women must work to be eligible for PL pay (indicated as "yes"); pay is offered regardless of work status prior to birth ("no"); and whether an alternative monthly fixed pay is offered for the otherwise non-eligible ("partial"). We then proceed by searching over the grid of possible parameter combinations, iterating at each parameter combination until the policy is revenue neutral.

	Max Welfare	Highest LFP	Possible Set
Paid leave duration	1 year	1 year	0,1,2,3
Job protection	2 year	1 year	0,1,2,3
Payment scheme	max($1000,55%)	max($500, 85%)	RR in 5% incr.
			Fixed mnth. pay in $500 incr.
Eligibility criteria	Partial	Yes	Yes/No/Partial
Pay for non-eligible	$500	$0	$0-$1500, in $250 incr.
Effect on fertility	0.3 children	0.3 children	N/A
Effect on LR-LFP	0.0	0.1	N/A

Table 3: Optimal PL policy parameters that maximize welfare and female labor force participation

The results in Table 3 show that both the welfare-maximizing and LFP-maximizing policy

[12]We do not model child health or child skill accumulation, and therefore we will not focus on this alternative benchmark. However, the value that parents place on child quality is captured indirectly in our model, through the increase in the utility after the childbirth from time spent not working in the labor market.

designs share many similarities. They offer paid leave for one year at relatively high replacement rate, and both require work prior to birth to establish eligibility for the full policy. The LFP maximizing policy has a higher replacement rate (85%), as this more strongly incentivizes labor supply. The reason the rate is lower than 100% is that the additional incentive effects from the higher replacement rates are small, but the cost of the policy rises, generating disincentives to work through the higher required taxation in other periods. The LFP-maximizing policy offers a floor to the PL pay to also incentivize the lowest wage workers to have at least some positive hours. Those who do not participate prior birth receive no PL payments, consistent with the objective of the planner to maximize work.

While the welfare-maximizing policy shares many similarities with the LFP-maximizing policy, the replacement rate is lower (55%) and the minimum pay higher ($1000), which lowers the regressiveness of the policy, and better insures the lowest-income women in the year after birth. For similar reasons, the policy also offers a minimum pay of $500 for one year even to women who were not working prior to birth. Since unexpected pregnancies are possible in our model, this transfer in particular helps those who had unexpected pregnancies early in the lifecycle when both income and wealth are low. Compared to the LFP-maximizing policy, the welfare-maximizing policy also allows for an additional year of unpaid leave, trading off work incentives for more flexibility to take time off during a period with relatively high disutility from work.

Nevertheless, the welfare-maximizing policy maintains relatively significant work incentives to reduce the efficiency losses from a paid leave policy. Consequently, the effect on long run labor force participation – ten years after first birth – compared to having no PL policy is approximately zero. Not surprisingly, the LFP-maximizing policy increases long-run LFP by a much larger ten percentage points. Interestingly, both policies have comparable (positive) effects on fertility. This indicates that some PL policy designs can have positive effects on fertility without reducing employment and, in some cases, even generating a positive effects on fertility and labor force participation simultaneously.

7 Conclusion

In this paper, we offer new evidence about the effects of different parental leave policy designs on women's labor supply and fertility decisions. We first show that most existing empirical studies of PL reforms – specifically those that exploit regression discontinuities around PL reform dates – offer exceptionally clean identification, but estimate parameters that are not easily interpretable and are potentially misleading to a policymaker. In particular, we show that RDD-style quasi-experiments are prone to underestimate – in some cases, severely – the effects of PL reforms. Consequently, we argue that this existing quasi-experimental evidence does not warrant the conclusion that PL reforms have “virtually no impact” on women’s outcomes (e.g., Kleven et al. 2021). We then proceed in three steps.

First, we document novel facts about women’s transitions between work, parental leave, and homemaker status after first birth. We focus on four countries with different parental leave policy designs: Sweden, Austria, Germany, and the UK. Policy designs in these countries vary primarily in leave duration, duration of job protection, and payment scheme. We show that transitions between these states in the ten years and “child penalties” after first birth differ significantly across the four countries in ways that are consistent with the different incentives embedded in their respective PL systems.

In the second step, we develop a dynamic structural model in which women make decisions about labor supply, consumption, and when and how many children to have, taking into account the effects of these decisions on human capital accumulation and depreciation, as well as on their parental leave benefits. We model the PL policy design in detail, along with childcare costs and subsidies and household income taxation. We calibrate the model using data from the UK BHPS.

In the third step, we conduct a series of counterfactual exercises to study in detail the effects of different PL policy designs. First, we vary the policy parameters to capture the different institutional environments in Sweden, Germany, and Austria. Simulating the model with these new policy parameters reproduces the main differences across these countries in labor supply and fertility.

Next, we simulate an environment in which two groups of women are exposed to identical “old” PL policies prior to their first birth; to identical ”new” policies for their second and third births; and who differ in their exposure to the new policy for their first birth only. This simulation replicates the RDD setting studied in the literature. We show, first, that the model replicates the finding that differences in labor supply and fertility decisions between the two groups are negligible. We also show that when the two groups of women are exposed to different policies over their full childbearing years, which is what we define as the ideal experiment, the differences in labor supply and fertility decisions are quantitatively significant.

Lastly, we derive the PL parameters that optimize welfare, female labor force participation, and fertility. We show that a policy that is short to moderate in length and that offers PL compensation based on wage replacement has positive effects on both fertility and labor supply.

References

Andresen, Martin Eckhoff and Emily Nix, "You Can't Force Me Into Caregiving: Paternity Leave and the Child Penalty," 2024.

Angelov, Nikolay, Per Johansson, and Erica Lindahl, "Parenthood and the gender gap in pay," *Journal of Labor Economics*, 2016, *34*.

Attanasio, Orazio, Peter Levell, Hamish Low, and Virginia Sánchez-Marcos, "Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply," *Econometrica*, 2018, *86*.

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

Baker, Michael, Jonathan Gruber, and Kevin Milligan, "Universal child care, matemal labor supply, and family well-being," *Journal of Political Economy*, 2008, *116*.

Bartel, Ann P., Maya Rossin-Slater, Christopher J. Ruhm, Jenna Stearns, and Jane Waldfogel, "Paid Family Leave, Fathers' Leave-Taking, and Leave-Sharing in Dual-Earner Households," *Journal of Policy Analysis and Management*, 2018, *37*.

Baum, Charles L., "The effect of state maternity leave legislation and the 1993 Family and Medical Leave Act on employment and wages," *Labour Economics*, 2003, *10*.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the gender gap for young professionals in the financial and corporate sectors," *American Economic Journal: Applied Economics*, 2010, *2*.

Bick, Alexander, "The Quantitative Role of Child Care for Female Labor Force Participation and Fertility," *Journal of the European Economic Association*, 6 2016, *14*, 639–668.

_ **and Nicola Fuchs-Schündeln**, "Taxation and labour supply of married couples across countries: A macroeconomic analysis," *Review of Economic Studies*, 2018, *85*.

Bronson, Mary Ann and Maurizio Mazzocco, "Taxation and Household Decisions: an Intertemporal Analysis *," 2022.

_ **and Peter Skogman Thoursie**, "The Wage Growth and Within-Firm Mobility of Men and Women: New Evidence and Theory," *Working Paper*, 2021.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes, "A flying start? Maternity leave benefits and long-run outcomes of children," *Journal of Political Economy*, 4 2015, *123*, 365–412.

Cools, Sara, Jon H. Fiva, and Lars J. Kirkebøen, "Causal Effects of Paternity Leave on Children and Parents," *Scandinavian Journal of Economics*, 2015, *117.*

Corekcioglu, Gozde, Marco Francesconi, and Astrid Kunze, "Expansions in Paid Parental Leave and Mothers' Economic Progress," *SSRN Electronic Journal*, 2022.

Dahl, Gordon B., Katrine V. Løken, Magne Mogstad, and Kari Vea Salvanes, "What is the case for paid maternity leave?," *Review of Economics and Statistics*, 10 2016, *98*, 655–670.

Ekberg, John, Rickard Eriksson, and Guido Friebel, "Parental leave - A policy evaluation of the Swedish "Daddy-Month" reform," *Journal of Public Economics*, 2013, *97.*

Farré, Lídia and Libertad González, "Does paternity leave reduce fertility?," *Journal of Public Economics*, 4 2019, *172*, 52–66.

Gayle, George-Levi and Andrew Shephard, "Optimal Taxation, Marriage, Home Production, and Family Labor Supply," *Econometrica*, 2019, *87.*

Goldin, Claudia, Sari Pekkala Kerr, and Claudia Olivetti, "Why Firms Offer Paid Parental Leave: An Exploratory Study," *NBER Working papers*, 2020.

Havnes, Tarjei and Magne Mogstad, "No child left behind: Subsidized child care and children's long-run outcomes," *American Economic Journal: Economic Policy*, 2011, *3.*

Keane, Michael P., "Labor supply and taxes: A survey," *Journal of Economic Literature*, 2011, *49.*

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 10 2019, *11*, 181–209.

_ , _ , Johanna Posch, Andreas Steinhauer, and Josef Zweimüller, "Child Penalties across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 5 2019, *109*, 122–126.

_ , _ , _ , _ , and _ , "Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation," *SSRN Electronic Journal*, 11 2021.

Lalive, Rafael and Josef Zweimüller, "How does parental leave affect fertility and return to work? Evidence from two natural experiments," *Quarterly Journal of Economics*, 2009, *124.*

Lefebvre, Pierre and Philip Merrigan, "Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada," *Journal of Labor Economics*, 2008, *26.*

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen, "Can women have children and a career? IV evidence from IVF treatments," *American Economic Review*, 2017, *107.*

Nyberg, Anita, *Lessons from The Swedish Experience*, Sydney: Sydney University Press,

OECD, "Starting Strong Early Childchood Education and Care Policy. Country Note for Austria," 2006.

_ , "Parental leave systems," 2021.

Olivetti, Claudia and Barbara Petrongolo, "The economic consequences of family policies: Lessons from a century of legislation in high-income countries," *Journal of Economic Perspectives*, 2017, *31*, 205–230.

Patnaik, Ankita, "Reserving time for daddy: The consequences of fathers' quotas," *Journal of Labor Economics*, 2019, *37.*

Raute, Anna, "Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits," *Journal of Public Economics*, 1 2019, *169*, 203–222.

Rossin-Slater, Maya, Christopher J. Ruhm, and Jane Waldfogel, "The effects of california's paid family leave program on mothers' leave-taking and subsequent labor market outcomes," *Journal of Policy Analysis and Management*, 2013, *32*, 224–245.

Ruhm, Christopher J., "The economic consequences of parental leave mandates: Lessons from Europe," *Quarterly Journal of Economics*, 1998, *113*, 285–317.

Schönberg, Uta and Johannes Ludsteck, "Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth," *Journal of Labor Economics*, 2014, *32*, 469–505.

Stearns, Jenna, "The Long-Run Effects of Wage Replacement and Job Protection: Evidence from Two Maternity Leave Reforms in Great Britain," *SSRN Electronic Journal*, 9 2017.

Online Appendix

for "Female Labor Supply, Fertility and Parental Leave Policy Design"

by Mary Ann Bronson and Deniz Sanin

A1 Additional Tables

Table A1: Summary Statistics for Women in Sweden, Austria, Germany and the UK

	Sweden		Austria		Germany		UK	
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
All Event Times								
LFP	0.93	0.00	0.69	0.01	0.61	0.00	0.77	0.01
Observations	3284		4000		23813		2844	
Event Times: -5 to -1								
LFP	0.97	0.00	0.94	0.01	0.90	0.00	0.97	0.01
Event Times: 3 to 5								
LFP	0.66	0.00	0.66	0.00	0.66	0.00	0.66	0.00
Event Times: 5 to 10								
LFP	0.87	0.02	0.72	0.01	0.65	0.00	0.76	0.01
By Event Time 10								
Minimum 2 Births	0.88	0.01	0.79	0.03	0.74	0.01	0.68	0.03
Minimum 3 Births	0.28	0.02	0.17	0.02	0.27	0.01	0.16	0.03

Notes: SEM stands for the standard error of the mean. The Sweden dataset consists of appending the two waves of the Sweden Generations and Gender Survey (GGS). For before first birth outcomes, we use the second wave of the Swedish GGS, which is compiled from Swedish social registers. For after first birth outcomes, we use the first wave of the Swedish GGS which is cross-section. Thus, one individual represents a specific event-time. For instance, if a mother gave her first birth in 2005, and the survey year is 2012 (can also be 2013), the event-time is 7 (7 years after first birth). We use women who gave birth after 1992. For the UK, both before and after birth outcomes are coming from British Household Panel Survey (BHPS). We restricted the sample to the first births between 1995 and 2000 due to the policy changes in the country. For Austria, both before and after birth outcomes are coming from the Wave 2 of Austrian GGS. The retrospective questions about women's statuses enable us to create a history of statuses for each event-time. We restrict to sample to the first births between 1996 and 2000 (1996 and 2000 included) due to the policy changes in the country. Lastly, both of before and after birth outcomes in Germany are coming from German Socio-Economic Panel (SOEP). We restricted the sample to the first births between 1993 and 2005 due to the policy changes in the country. We keep the women born after 1960 in the whole sample to approximate a balanced number of women for each country.

A2 Additional Figures

A2.1 Additional Figures for Section 5: Results

Figure A1: Women' s Activity Status Shares within Each Country

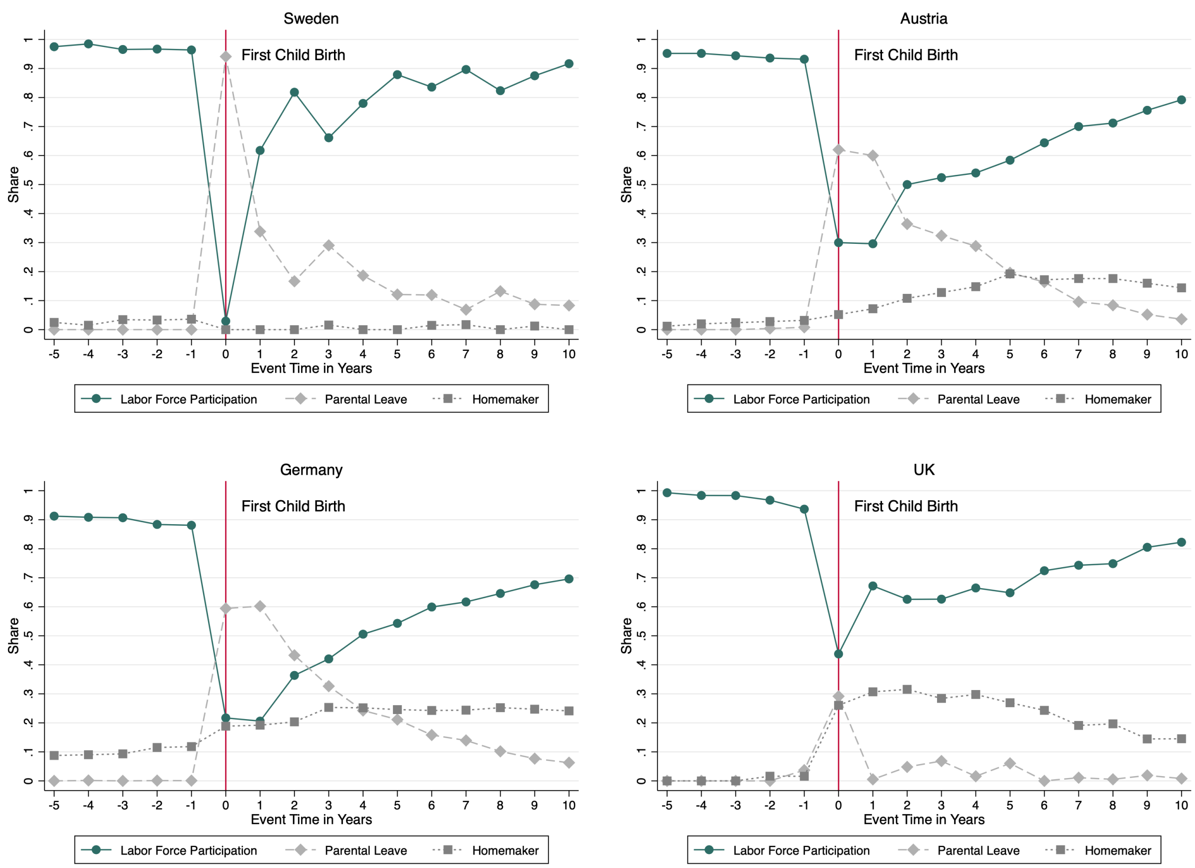


Figure A2: Women's Labor Market and Fertility Outcomes in Sweden and UK, by Years Relative to First Birth (with control variables)

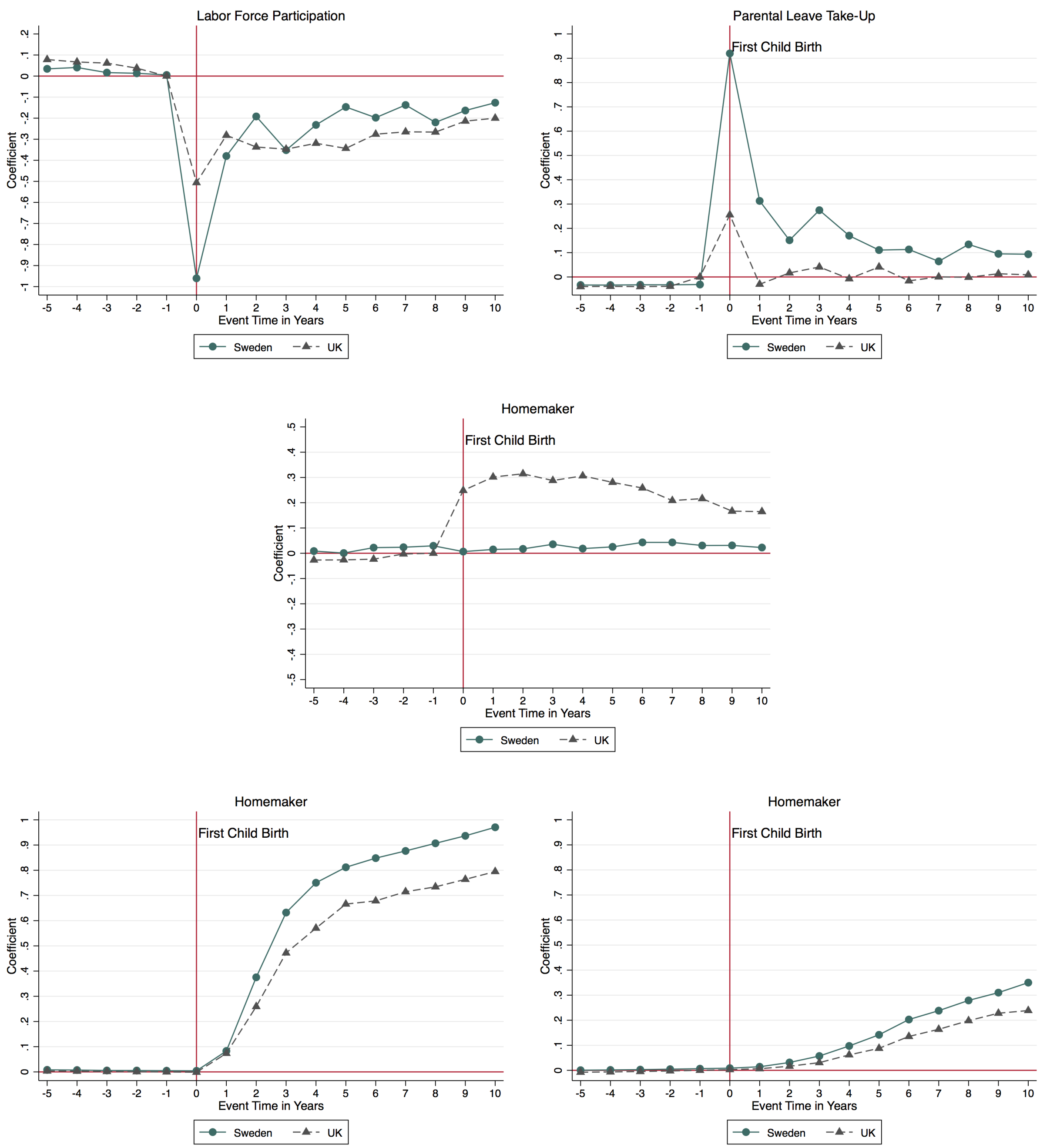


Figure A3: Women's Labor Market and Fertility Outcomes across Parental Leave Policy Designs, by Years Relative to First Birth (with control variables)

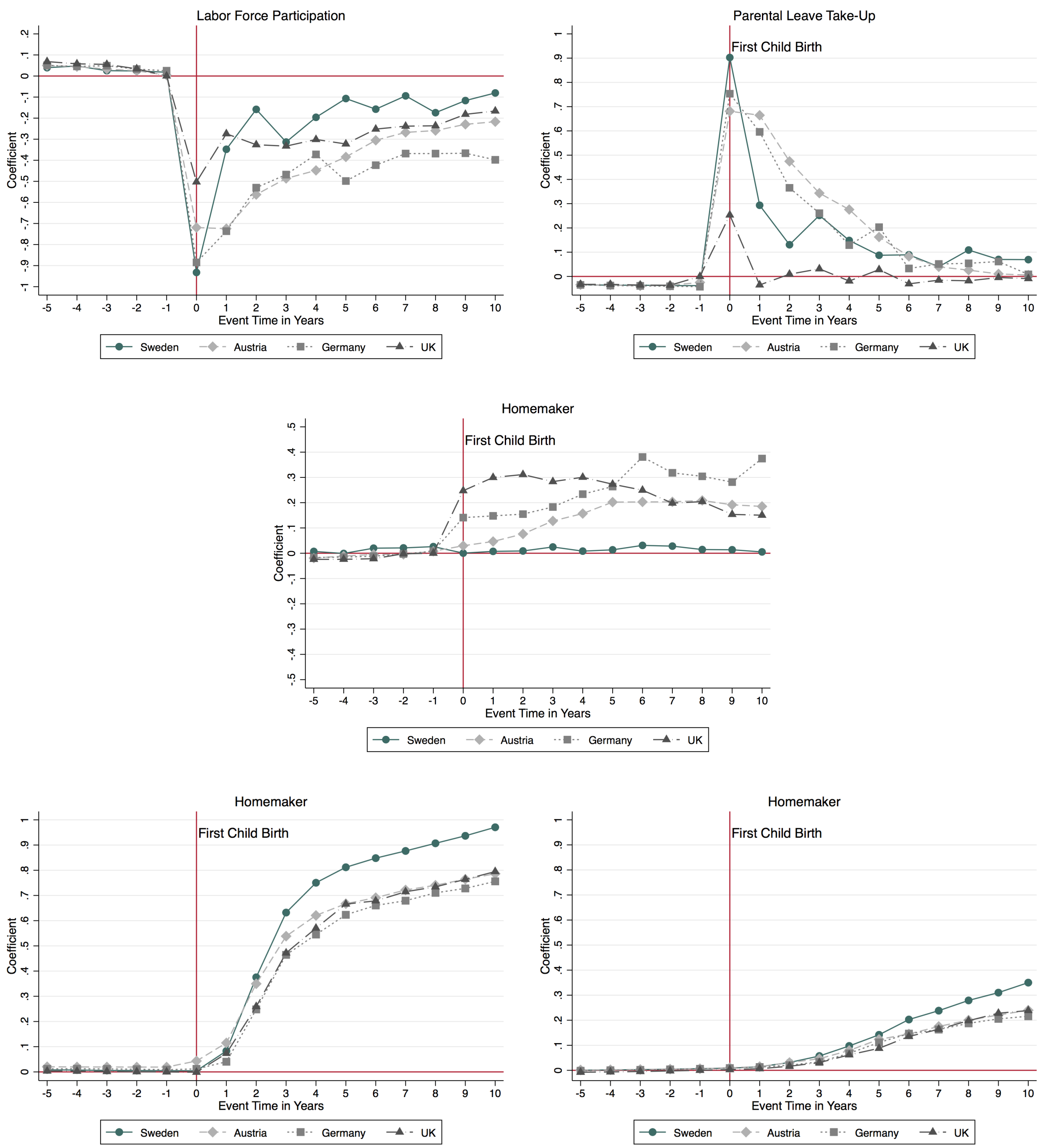


Figure A4: Men's Labor Market Outcomes across Parental Leave Policy Designs, by Years Relative to First Birth

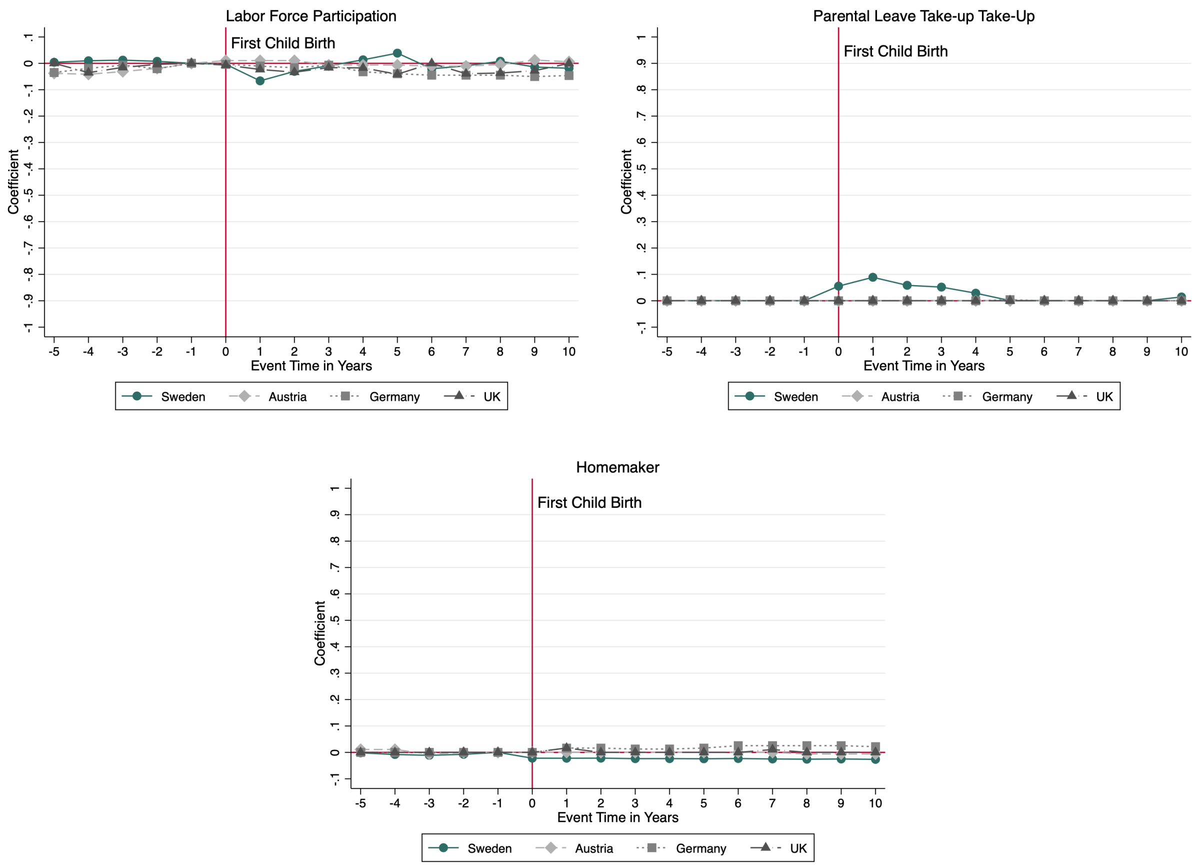


Figure A5: Men's Labor Market Outcomes across Parental Leave Policy Designs, by Years Relative to First Birth (with controls)

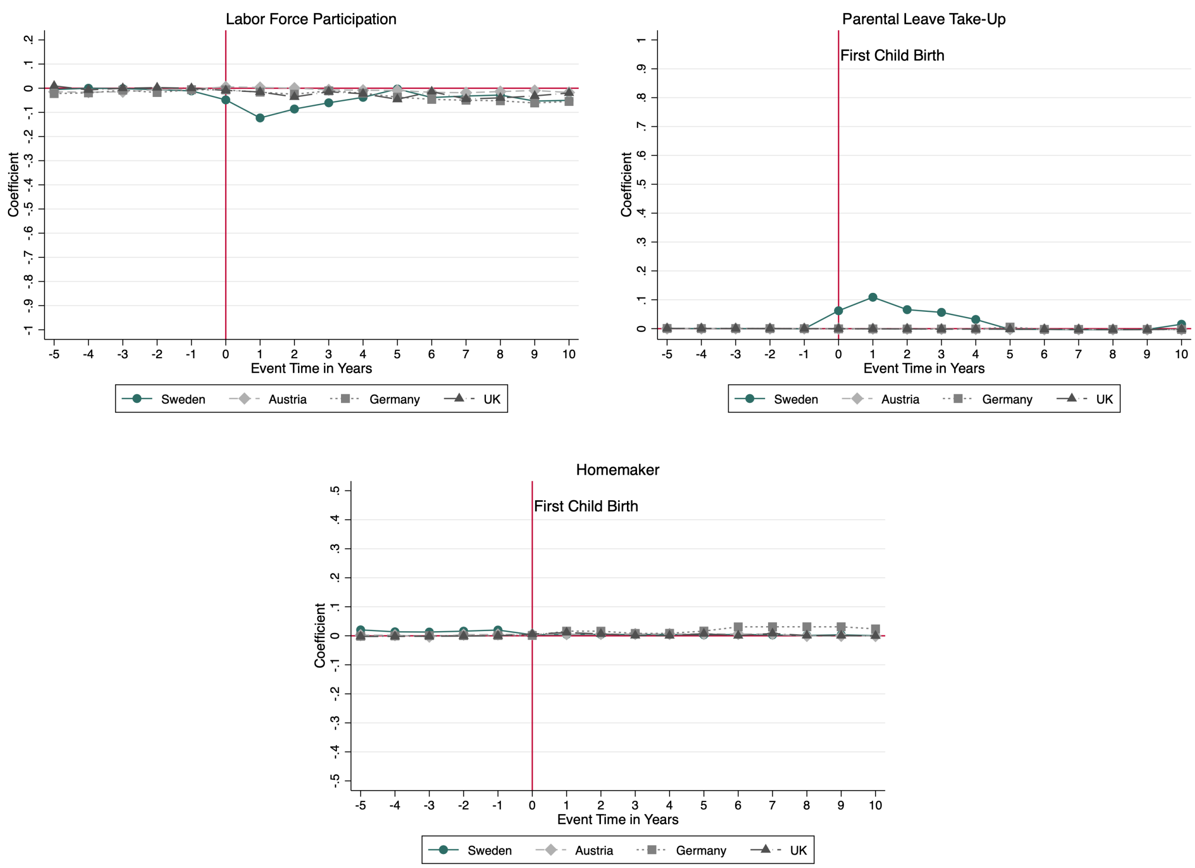


Figure A6: Share of Women on Parental Leave, by Number of Children

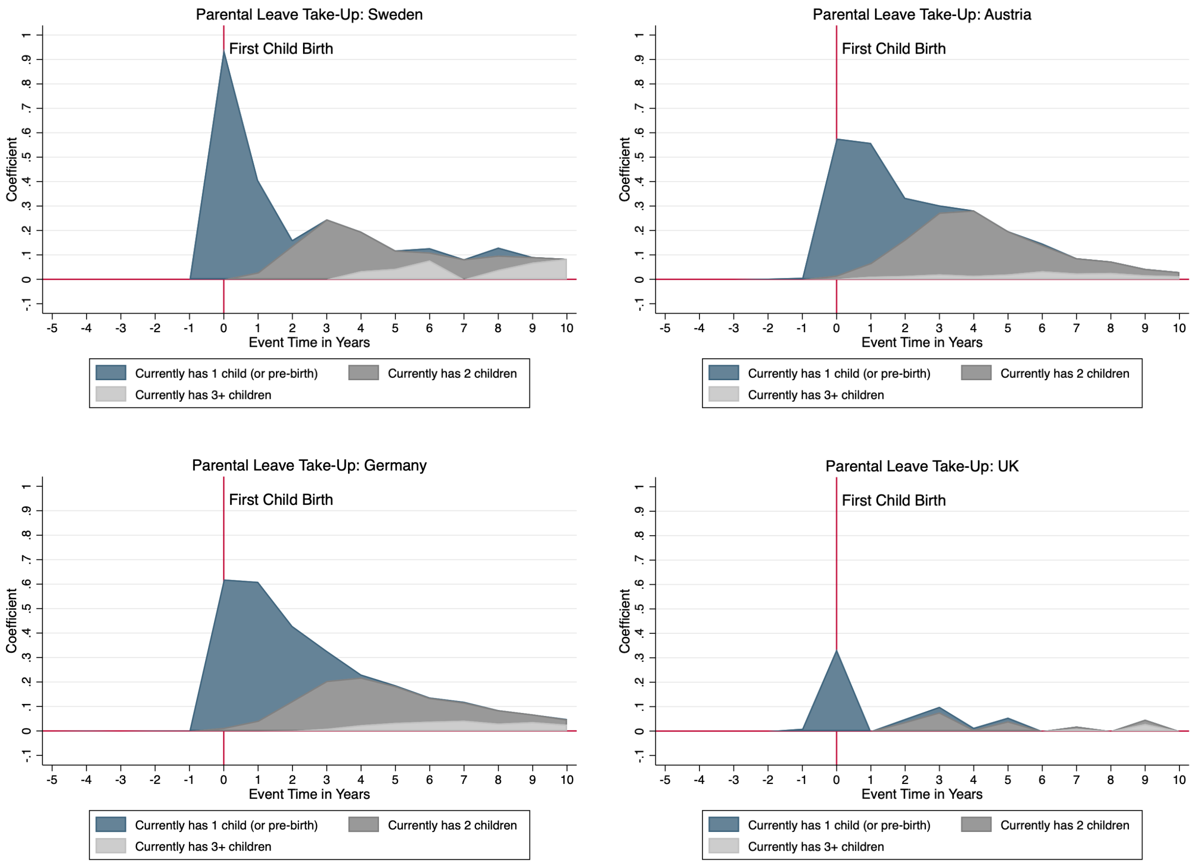
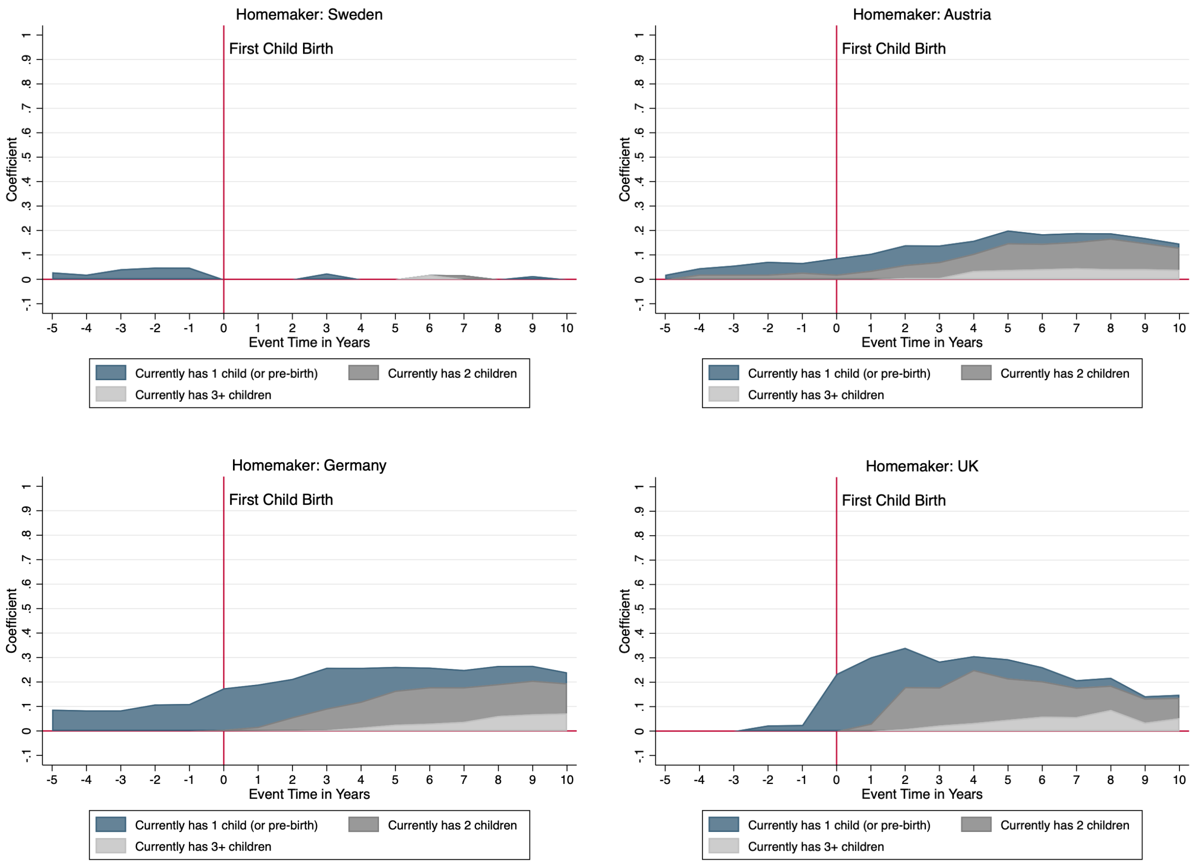


Figure A7: Share of Women Who are Homemakers, by Number of Children



A3 Policy Appendix

A3.1 Parental Leave Policies

A3.1.1 Austria

- **1974:** 16 weeks of mandatory job-protected maternity leave is introduced. 100% of earnings are replaced.
- **1996:** Maximum amount of parental leave a mother can take is 78 weeks. 26 weeks of father quota is introduced.
- **2002:** All citizens are eligible for parental leave without an employment requirement. Maximum amount of parental leave a mother can take is 130 weeks. Cash-for-care benefits and parental leave entitlements were separated. If both parent apply for cash-for-care benefit and participate in the labor market for at least 26 weeks, parents receive the childcare benefits at a flat rate for 156 weeks.

A3.1.2 Germany

- **1968:** Mandatory paid maternity leave is introduced. 100% of earnings are replaced.
- **1992:** Total parental leave is for 148 weeks where 70 weeks of the leave is paid and the remaining 78 weeks are unpaid. Job-protected unpaid parental leave can be used until the child's 3rd birthday.
- **1993:** Duration of paid parental leave is extended to 96 weeks. Unpaid parental leave is for 52 weeks.
- **2001:** Options for paid parental leave is introduced. Mothers have the option to take a shorter and better paid leave (DM 900 for 44 weeks) compared to a longer (DM 600 for 96 weeks) but less well paid leave. Duration of unpaid leave is unchanged. Paid leave can be taken until the child's second birthday.

A3.1.3 Sweden

- **1955:** Sweden introduced its first paid maternity leave policy.
- **1974:** Maternity leave is replaced by parental leave. The mother and the father could share 26 weeks of job-protected parental leave. 80% of earnings are replaced.
- **1994:** The maximum amount of job-protected paid parental leave mothers can take was 52 weeks which is paid for 80% of the parent salary.
- **1995:** Parental leave was divided in to mommy and daddy months. Parents could give months to each other except for one month which were called the mommy and daddy month. Sweden is the second country in the world that introduced a daddy month, a one month father quota for parental leave. 90% of earnings were replaced during the momy and daddy month. In the remaining months, 80% of earnings were replaced. 13 weeks of additional parental leave is introduced which is paid at a flat rate.
- **1996:** Replacement rate was reduced to 75% for the non-daddy and non-mommy months. Replacement rate for the mommy and the daddy month was reduced to 85%.
- **1997:** Replacement rate for the mommy and the daddy month were reduced to 75%.
- **1998:** Replacement rate was raised to 80% for the mommy and the daddy months and for 43 weeks of the parental leave. The remaning weeks were paid at a flat rate.
- **2002:** Paid parental leave increased to 69 weeks of which 56 weeks are paid for 80% of the parent salary and the remainig 13 weeks are paid at a flat rate. The father quota increased to 2 months.

A3.1.4 UK

- **1994:** Mothers who have been with their employers for 26 consecutive weeks were entitled to 18 weeks of job-protected maternity leave, where the first 6 weeks were paid for 90% of the mother's salary and the remaining 12 weeks were paid at a flat rate.
- **1999:** Fathers and mothers could take upto 13 weeks of unpaid parental leave each.
- **2000:** All employed pregnant women were entitled to 18 weeks of job-protected maternity leave.
- **2003:** Maternity leave was extended to 52 weeks. The first 6 weeks are paid for 90% of the mother's salary. Any 20 weeks following the first 6 weeks were paid at a flat rate. The

remaining 26 weeks were unpaid. In addition to the maternity leave, mothers and fathers could still take 13 weeks of unpaid parental leave.

- **2007:**[1] The flat rate payment period was extended to 33 weeks. Maternity leave duration remained as 52 weeks. The first 6 weeks are paid for 90% of the mother's salary. Any 33 weeks following the first 6 weeks were paid at a flat rate. The remaining 13 weeks are unpaid. In addition to the maternity leave, mothers and fathers could still take 13 weeks of unpaid parental leave. Father could take 2 weeks of paternity leave during the first 8 weeks following childbirth and it was paid at a flat rate.

A3.2 Paternity Leave Policy

Within our sample period, in 1995, Sweden introduced a "daddy month," a one month father quota for parental leave. The quota reserves a part of the parental leave period specifically to fathers. If the father does not take the leave, the family loses the reserved leave period. Similar to Sweden's parental leave policy, Swedish daddy month is also "high-paid".[2] The quota increased to 2 months in 2002. Although Austria introduced a 6 month father quota during our sample period, it is paid at a flat rate. Thus, it is a "low-paid" father quota. Germany introduced 2 months of father quota at the very end of our sample period (2007) and the UK did not introduce any.

A3.3 Childcare Policies

Among the four countries we consider, Sweden also has the most generous childcare policies. Germany, Austria and the UK are similar, and offer public pre-school starting at age 3. Consequently, the difference between the age of the child by the end of the paid parental leave and the child's child care starting age is the smallest in Sweden.[3]

During our sample period, access to public or subsidized child care under age 3 was very limited in Germany, the UK and Austria. Public child care availability becomes more prevalent after age 3. In Germany, in the early 2000s, subsidized child care slots are available only to 6% of children aged 0-2 (Bick, 2016). Children aged 3-6 go to pre-school (Kindergarten). Subsidized slots are available to 96% of the children in this age group. In the UK, during the early 1990s, the provision of child care is mostly linked to issues of unemployment and welfare dependency. In

[1] This change is not used in the paper since only 8.5% of the first births in the UK sample took place in 2007 and 2008.

[2] "High paid" term is used as in the language of Figure 2 to highlight that the father's salary is replaced with at least 80%.

[3] In 1995, a Swedish law came into effect that pushed municipalities to provide public child care for children over 18 months. In 1996, 45% and 69% of one and two year-olds respectively were in public child care in Sweden (Nyberg, 2007).

1998, the government launched a program to provide nursery education for all 3 years-old children by 2004 (?). In Austria, publicly provided child care services are divided into two. “Krippen” is centre-based creche services for children under age 3 and “Kindergarten” is for children aged 3-6. Although Austria provides childcare for children aged less than 3 years old, only 5% of the age cohort attend these centers (84% of the age cohort attend kindergartens) (OECD, 2006). This is related to the fact that crèches are found most often in large urban areas. Apart from the capital, coverage by crèche services is low and reaches less than 5% of 0-3 aged children across the rest of the country (OECD, 2006). Thus, for majority of Austrian mothers, like in Germany and the UK, public child care for children under age 3 is not provided.

Relationship between parental leave and child care policies. Due to the differences in parental leave policies across these countries, the difference between the age of the child by the end of the paid parental leave and the child’s child care starting age varies. The differences are visualized in Figure A8. Since paid parental leave is 69 weeks in Sweden, the difference between the age of the child by the end of the paid parental leave (low paid) and the child’s child care starting age is 3 weeks. The difference is 16 weeks if we substract the age of the child by the end of the high paid leave from the child care starting age. These differences are much higher in Austria and Germany, and highest in the UK.

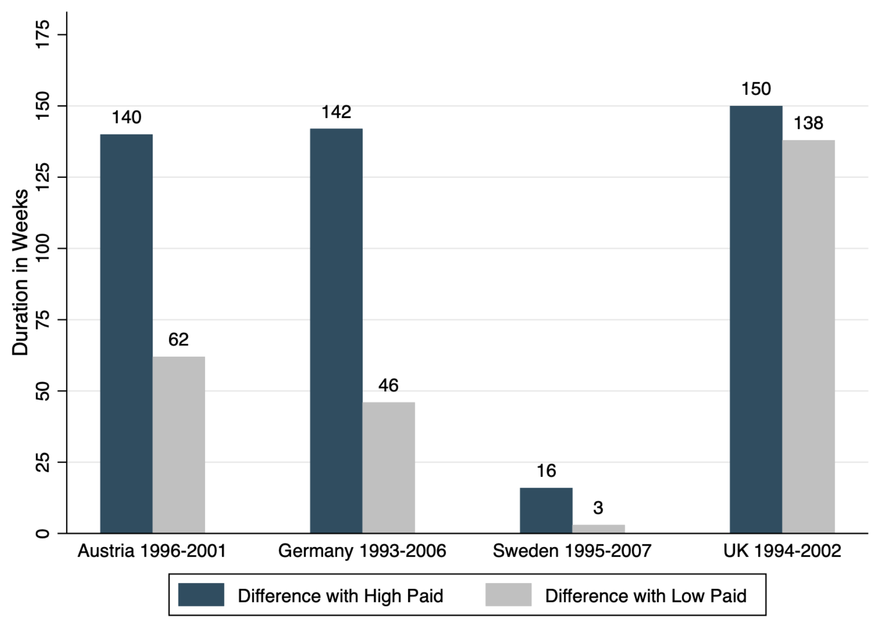


Figure A8: The difference between the age of the child by the end of the paid parental leave and the child’s child care starting age in Europe

Notes: The difference is calculated by (the child’s child care starting age)-(the age of the child by the end of the high/low paid parental leave)

A4 Data Appendix

We combine three, nationally representative data sources for our analysis. First, we use a unique dataset, Generations and Gender Survey (GGS), which is a panel survey of adults in mostly European 19 countries. It is conducted in two waves during the 2000s. We use Austria and Sweden surveys for our analysis. The second wave of Swedish GGS was entirely collected from administrative registers.

GGS collects rich demographic information on labor supply and fertility. In Austrian survey, the information on labor supply is collected retrospectively during the interview: individuals are asked to report each of their activity statuses and their start and end dates since age 16. All statuses are mutually exclusive. Based on the retrospective information, we build an individual/year panel. We define activity status of a respondent in a given year as their reported activity status. An individual is defined as in labor force if she reported being either employed, self-employed, at school or unemployed in a year. Similarly, an individual is defined as on parental leave if she reported being on maternity or parental leave. Lastly, a respondent is defined as a homemaker, if the respondent reported being an homemaker in a year. For information on fertility, surveys ask women to report their complete birth histories. Based on the histories provided, we construct individuals' number of children in each year.

Wave 2 of Swedish GGS collects information on respondents from administrative registers for every year between 1992-2015, where Wave 1 have rich survey data on respondents' activity statuses and birth history. We use Wave 1 for event times after first birth and Wave 2 for event times before first birth. This enables us to have consistent (and granular) information on activity statuses across countries. When we are constructing variables for Sweden for event times bigger than zero, using the survey data, we follow the following rules. An individual is defined as in labor force if she reported being either employed, self-employed, at school or unemployed in a year. Similarly, an individual is defined as on parental leave if she reported being on maternity or parental leave. Lastly, a respondent is defined as a homemaker, if the respondent reported being an homemaker in a year. For information on fertility, surveys ask women to report their complete birth histories. Based on the histories provided, we construct individuals' number of children in each year. Wave 2 is a cross-section dataset. Thus, one individual represents a specific event-time. For instance, if a mother gave her first birth in 2005, and the survey year is 2012 (can also be 2013), the event-time is 7 (7 years after first birth).

When we are constructing variables for Sweden for event times smaller than zero, using the administrative registers, we follow the following rules. All statuses are mutually exclusive. An individual is defined as being on parental leave in a given year if the individual received parental leave benefit more than 91 days in that year. An individual is defined as employed in a given year if she has received a positive amount payment from a main job or business in that year and she is not on parental leave. Being a student as an activity status is defined as participating in the autumn term of that year and not being employed or on parental leave. Unemployment is defined as received a positive amount of unemployment benefit and not being employed, a student or on parental leave. As for other countries, we define an individual as in labor force if the individual is either employed, student or unemployed in a year. Lastly, being a homemaker is defined as not being any of the statuses above. This means that homemakers do not receive an income and they are not students.

We use British Household Panel Survey (BHPS) to study the UK. BHPS is a panel survey of adults in the UK which interviews individuals every year from 1991 to 2008. The survey asks respondents to report their current labor force status in each wave. We focus on the statuses we use in Austrian, German and Swedish GGS: employed, self-employed, parental leave or family care. All statuses are again mutually exclusive. We again define an individual as employed if she reported being either employed or self-employed in a year. Similarly, an individual is defined as on parental leave if she reported being on maternal leave. Lastly, a respondent is defined as a homemaker, if the respondent reported family care as her labor force status in a year. The survey also asks women for their birth histories. Based on the information, we again construct the number of children of individuals each year.

Lastly, we use The German Socio-Economic Panel (SOEP) to study Germany. SOEP is a panel survey of adults in Germany which interviews individuals every year between 1984 to 2021. As in BHPS, the survey asks respondents to report their current labor force status and birth histories in each wave. We create variables on labor force status and fertility based on the procedure we follow when we are creating variables in BHPS.

We combine the data for all countries and restrict the sample to individuals who have at least one children. For Austria, the years of first birth are between 1996-2000. For UK, the years of first births are between 1995 and 2000.[4] For Sweden, we use women who gave birth between 1992 and 2008.

[4]These samples are based on policy changes in the respective countries.

A5 Counterfactual Policy Experiments Appendix

A5.1 Baseline Counterfactuals: Simulating Germany

Germany and Austria have fairly similar PL and childcare policies. However, according to the simulation results, German mothers have a much lower labor supply. The model generates this due to two reasons. First, there are differences in taxation: In Germany's joint tax system, secondary earners have high marginal tax rates that disincentivize work. Women's lower labor supply generates negative income and wealth effects, which reduce fertility rates compared to Austria.

Second, recall that German mothers can extend their leave one more year (they can take up to three years of PL) without receiving any pay during this extended period. If women's wage growth is based on women's human capital accumulation (how much they worked and gained experience in their careers), German mothers' wage growth after first birth is plausibly slower compared to Austrian mothers. This is related to German mothers taking more time off their work due to the longer duration of the PL policy in Germany. More time out of the labor force after first and subsequent births may result in lower human capital accumulation and earning dynamics after the first birth compared to Austria. Low earnings can disincentivize German mothers to return to the labor force after subsequent births.

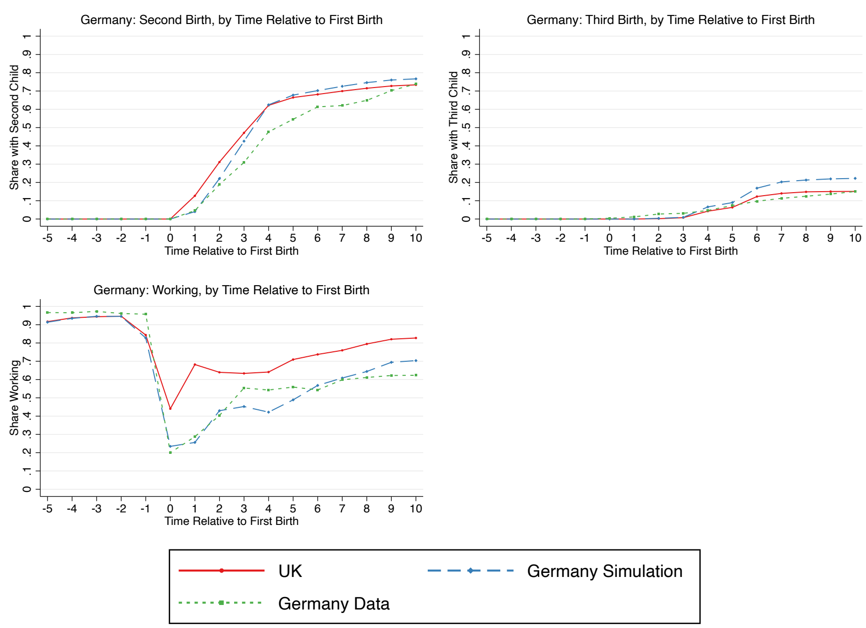


Figure A9: Baseline Counterfactual: Germany