Child Penalties Across Countries:
Evidence and Explanations

Online Appendix

Henrik Kleven, Princeton University and NBER

Camille Landais, London School of Economics

Johanna Posch, European University Institute

Andreas Steinhauer, University of Edinburgh

Josef Zweimüller, University of Zurich

A   Supplementary Figures

Figure A.I: Child Penalties in Participation Rates in Scandinavian Countries
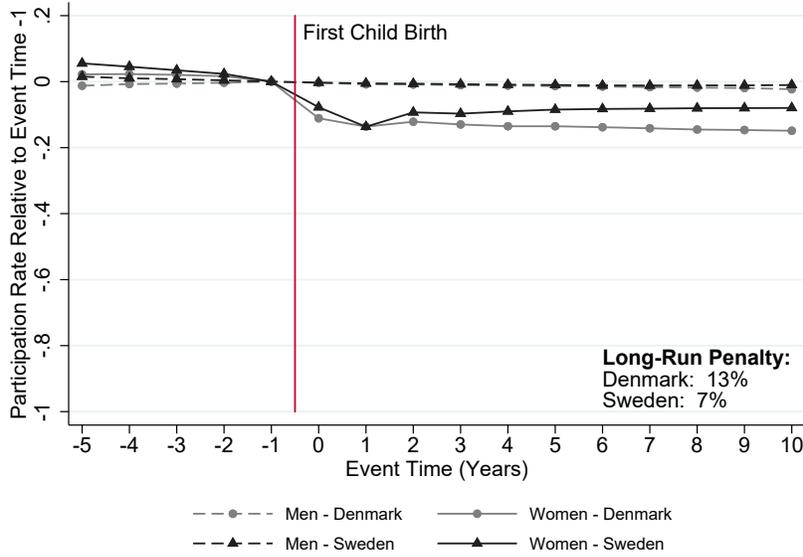
Notes: The figure shows percentage effects of parenthood on the participation rate across event time t for each gender g, i.e. P_t^g defined above. The figure also displays long-run child penalties, defined as the average penalty P_t from event time 5 to 10.

Figure A.II: Child Penalties in Participation Rates in English-Speaking Countries
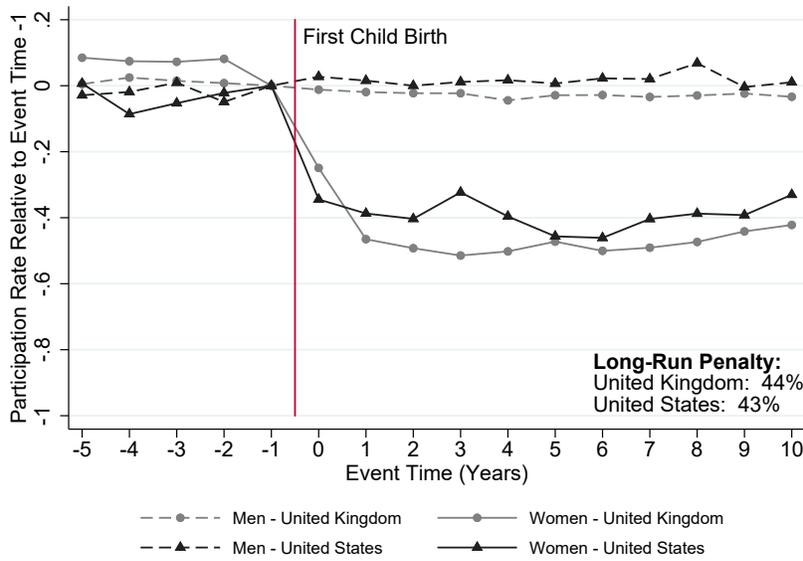
Notes: The figure shows percentage effects of parenthood on the participation rate across event time t for each gender g, i.e. P_t^g defined above. The figure also displays long-run child penalties, defined as the average penalty P_t from event time 5 to 10.

Figure A.III: Child Penalties in Participation Rates in German-Speaking Countries
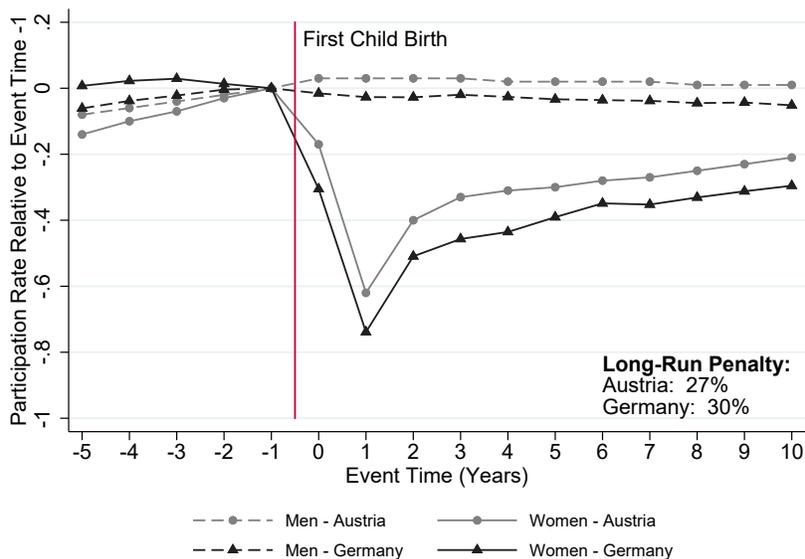
Notes: The figure shows percentage effects of parenthood on the participation rate across event time t for each gender g, i.e. P_t^g defined above. The figure also displays long-run child penalties, defined as the average penalty P_t from event time 5 to 10.

B   Data Description

The methodology applied to compute child penalties in this paper relies on Kleven *et al.* (2018). We estimate the impact of children on the labor market trajectories of mothers and fathers using event studies around the birth of the first child. This approach requires high-quality panel data with information on labor market outcomes and children.

In this section, we explain the data sources and sample selection criteria used for all 6 countries. In order to make the estimates as comparable as possible, we have followed the same sample selection principles in all the countries.

We consider individuals who have their first child between the ages of 20 and 45, and who are observed between 5 years before and 10 years after child birth. Our analysis focuses on first child births where the parents are known and alive. We do not impose any restriction on the relationship status of parents and include all individuals who have a child in a given year and can be followed over time whether or not they are married, cohabiting, separated, divorced, or have not yet formed a couple in any given year.

B.1 Austria

B.1.1 Data Sources

The Austrian dataset consists in the Austrian Central Social Security Register (ASSD), which provides very detailed longitudinal information and covers the whole population of dependent employees between 1980 and 2017.

Data Availability   The original ASSD data can be obtained from Statistics Austria. To that end, please contact Statistics Austria information services.

B.1.2 Variable Definition

Earnings   Yearly earnings are defined as the sum of the basic wage and supplementary payments such as bonuses, 13th and 14th monthly wages and extra vacation payments. Due to the fact that the lower tail of the earnings distribution is not observed in the original dataset, we are in fact ignoring the earnings of minor employment. Note that earnings in the social security data are top coded at the top of the earnings distribution. Earnings are adjusted to take inflation into account.

Employment   The ASSD records all employment spells for the universe of individuals. An individual is considered as working in a given year if there has been at least one employment spell in that year lasting one day or more.

Birth Information   Child births for mothers are registered in the social security data, but fathers are not observed in this dataset. Fathers are registered alongside mothers in the dataset of child benefits and matched to our main dataset until 2013. Event time is defined as the difference between the year in which earnings and employment status are reported and the date of birth of the first child. Due to the fact that civil servants are included in the sample but only observed after 1988, we might potentially misattribute some births as the first child when in fact they correspond to the second or third child. We conducted a robustness check excluding all individuals ever observed in public employment, making up 4% of the sample, and the results remained unchanged.

Table A.I: Descriptive Statistics

| Country | Year of First Child | | Age at First Child | | Number of Children |
	Range	Average	Men	Women	At $t = 10$
Austria	1985-2007	1995	30.0	26.5	1.7
Denmark	1985-2003	1994	28.5	26.2	2.2
Germany	1989-2005	1997	30.4	27.7	1.9
Sweden	1997-2011	2004	30.8	28.7	2.2
United Kingdom	1991-2008	1998	31.3	30.0	2.0
United States	1967-2006	1985	25.8	24.9	2.1

B.1.3 Sample

The sample contains all first child births where the parents are observed every year between 5 years before having a child and 10 years after. Descriptive statistics are reported in Table A.I.

B.2 Denmark

B.2.1 Data Sources

The analysis of child penalties in Denmark relies on administrative data for the full Danish population between 1980 and 2013. We combine several administrative registers linked at the individual level via personal identification numbers to obtain a dataset of parents containing their labor supply, demographic information and the date of birth of their first child. Crucially, the data allows us to link family members.

Data Availability   The data from our project can be requested from the Statistics Denmark's Division of Research Services (ask for a copy of our project #70629).

B.2.2 Variable Definition

Earnings   The Income Statistics Register contains information about all economically active individuals. Earnings are defined as labor income before taxes and transfers. They are observed at the calendar year level, i.e. as the sum of total gross earnings from January to December of each calendar year.

Employment   To define our employment status variable, individuals with reported non-zero earnings are considered as working.

Birth Information   Information on births stems from birth registry data and enables to link parents to their children. Event time is defined as the difference between the year relevant for earnings and the year of birth of the first child.

B.2.3   Sample

Our analysis focuses on first child births where the parents are known, alive and reside in Denmark. The sample contains all first child births where the parents are observed every year between 5 years before having a child and 10 years after. Descriptive statistics are reported in Table A.I.

B.3   Germany

B.3.1   Data Sources

Our German dataset is a representative longitudinal survey, the German Socio-Economic Panel (SOEP), which provides rich information on labor supply dynamics and allows to observe the date of birth of children to adult members of the sample. It was conducted between 1984 and 2016 in West Germany and between 1990 and 2016 in East Germany.

Data Availability   The data can be requested from the Research Data Center of the German SOEP.

B.3.2   Variable Definition

Earnings   When surveyed, individuals are asked to report the total of their gross labor income in the last 12 months. If an individual is surveyed in event year t, we therefore define their reported gross labor income in the survey as their earnings for event year $t-1$.

Employment   Employment status of the parent is defined based on wages and hours worked. More precisely, an individual is considered employed if he had positive wages and worked at least 52 hours in the previous year.

Birth Information Year of birth is reported for each child. Event time is defined as the difference between the survey year and the year of birth.

B.3.3 Sample

Our analysis focuses on first child births for the parents observed in the SOEP. The sample contains all first child births where the parents are observed every year between 5 years before having a child and 10 years after. Descriptive statistics are reported in Table A.I.

B.4 Sweden

B.4.1 Data Sources

The dataset for Sweden is a combination of fiscal statistics and registry data providing comprehensive information about the universe of births between 1997 and 2011.

Data Availability The data can be requested from Statistics Sweden MONA (Microdata Online Access) support.

B.4.2 Variable Definition

Earnings Information about earnings and employment status stems from the longitudinal integration database for health insurance and labour market studies (LISA), which includes all individuals who are at least 16 years old and registered in Sweden. Earnings are defined as labor income before taxes and transfers. They are observed at the calendar year level, i.e. as the sum of total gross earnings from January to December of each calendar year.

Employment To define the employment status variable, all individuals with non-zero earnings are considered as working.

Birth Information The birth dates of first children are inferred from the Swedish birth registry. Event time is defined as the difference between the year for which earnings and employment status are reported and the year of birth of the first child.

B.4.3 Sample

Our analysis focuses on first child births where the parents are known, alive and reside in Sweden. The sample contains all first child births where the parents are observed every year between 5 years before having a child and 10 years after. Descriptive statistics are reported in Table A.I.

B.5 United Kingdom

B.5.1 Data Sources

Our analysis of child penalties in the United Kingdom is based on the British Household Panel Survey (BHPS), following households for 18 years between 1991 and 2009.

Data Availability The BHPS can be obtained by contacting the UK data service.

B.5.2 Variable Definition

Earnings Reported earnings are defined as annual labor income before taxes and transfers.

Employment Participation is defined as having positive earnings last year.

Birth Information To identify parents who have had children before the BHPS starts, we use on fertility history interviews, conducted in the second wave after entry into the sample. Unfortunately, this fertility history is not asked again after the second wave, so we use the household grid, which lists all individuals who live in a household together, relationships between household members, and some basic demographic information, such as date of birth. In order to calculate event time, we calculate the difference in months between the dates of interview and of first child, divide it by 12, and round it to the nearest integer. Event times greater (less) than $t = 10$ ($t = -5$) are capped $t = 10$ ($t = -5$).

B.5.3 Sample

The sample focuses on first child births for the parents surveyed in the BHPS. It includes all individuals who have a child in a given year and can be followed over time. The restriction imposed

on the Danish administrative dataset in Kleven *et al.* (2018) that parents need to be observed every year would be excessively stringent in the case of the survey-based data in the United Kingdom. In order to mitigate selection effects and to ensure that the sample is a close as possible to a balanced panel with respect to event time, our sample includes parents who are observed (1) At least once before (5 years before first child birth to the year of birth) and once after their first child birth (one year after to 10 years after first child birth); (2) At least 8 times over the 15-year window. We further restrict age at first birth to be between 20 and 45. Descriptive statistics are reported in Table A.I.

B.6 United States

B.6.1 Data Sources

The Panel Study of Income Dynamics (PSID) follows a nationally representative sample of the American population from 1968 to present. Our dataset includes household heads and their spouse or cohabitant, for which we observe labour market outcomes and family history consistently.

Data Availability   The PSID data are available online and free of cost to researchers.

B.6.2 Variable Definition

Earnings   Given the objective of the PSID, information about earnings is precise and detailed. Earnings are defined as total labour income before taxes and transfers, including farm income, business income, wages, bonuses, overtime pay, commissions, as well as income from professional practice and roamers and boarders. They are reported by the respondents and their spouses or cohabitants for the year prior to the interview and are attributed to our earnings variable in year $t-1$. For individuals who are not working, earnings are coded as 0 rather than as missing values. There is censoring of reported income at the very top of the distribution; however, the thresholds are high and the number of observations affected by top-coding is tiny.

Employment   Our employment status variable is based on the answer to the question "Are you working now, unemployed, retired, or what?". We consider as working all the individuals who

indicate being working at the time of the survey, temporarily laid off or on sick leave, whereas unemployed, retired, permanently disabled individuals, students and individuals keeping house are considered as not working.

Birth Information The PSID asks about the birth year and month of the individual's first or only child. Date of birth of the first child is known with different degrees of precision, up to the month, the season or the year. In the two latter cases, several assumptions are made. First, we assume that children born in winter are born in January, and those born in the spring, summer or fall respectively in April, July and October. When first child birth is only known up to the year, we assume that it occurred in September. Event time is defined as the difference between the interview date and the birth date of the first child, rounded to the closest year. Interview date is available up to the fortnight from 1968 to 1979 and up to the day from 1980 to 2015, which allows to impute with relative precision the month of survey response.

B.6.3 Sample

The sample focuses on first child births where the parents are known, alive and reside in the United States. No restriction on the relationship status of parents is imposed and the sample includes all individuals who have a child in a given year and can be followed over time. The restriction imposed on the Danish administrative dataset in Kleven *et al.* (2018) that parents need to be observed every year would be excessively stringent in the case of the survey-based data in the United States. In order to mitigate selection effects and to ensure that the sample is a close as possible to a balanced panel with respect to event time, our sample includes parents who are observed (1) At least once before (5 years before first child birth to the year of birth) and once after their first child birth (one year after to 10 years after first child birth); (2) At least 8 times over the 15-year window. We further restrict age at first birth to be between 20 and 45. Descriptive statistics are reported in Table A.I.

B.7 Data on Gender Norms

B.7.1 Data sources

Data on gender norms stem from the International Social Survey Programme (ISSP). The relevant survey years are discussed in the following sections.

Data Availability The ISSP data are partly available online and upon request.

B.7.2 Data Treatment

Our study focuses on two survey questions: "Do you think that women should work outside the home full-time, part-time or not at all under the following circumstances?"

- "When there is a child under school age"

- "After the youngest child starts school"

For each country c and wave t, we define N_{ct} as the average fraction of individuals answering "should stay at home", across these two questions, weighting by survey weights. We then take for each country c the unweighted average of N_{ct} across all available waves. We consider all respondents aged between 16 and 64, regardless of their employment status and family situation. Moreover, the ISSP was conducted separately for West and East Germany, Northern Ireland and Great Britain. To compute the country-wide response rates, regions were weighted by their share in total population of the country. More precisely, we assumed that Northern Ireland and Great Britain make up 3% and 97% of population in the United Kingdom respectively, and West and East Germany 82% and 18% of total population in Germany respectively.

Table A.II: Available ISSP Waves for Gender-Related Questions By Country

| | **ISSP Wave** | | | |
Country	1988	1994	2002	2012
Austria	Yes	Yes	Yes	Yes
Denmark	No	No	Yes	Yes
Germany	Yes	Yes	Yes	Yes
Sweden	No	Yes	Yes	Yes
United Kingdom	Yes	Yes	Yes	Yes
United States	Yes	Yes	Yes	Yes

References

KLEVEN, HENRIK, LANDAIS, CAMILLE, & SØGAARD, JAKOB EGHOLT. 2018. Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, forthcoming.