# Online Appendix:

# "The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach"

Henrik Kleven
Princeton University and NBER

November 2025

# A Supplementary Exhibits

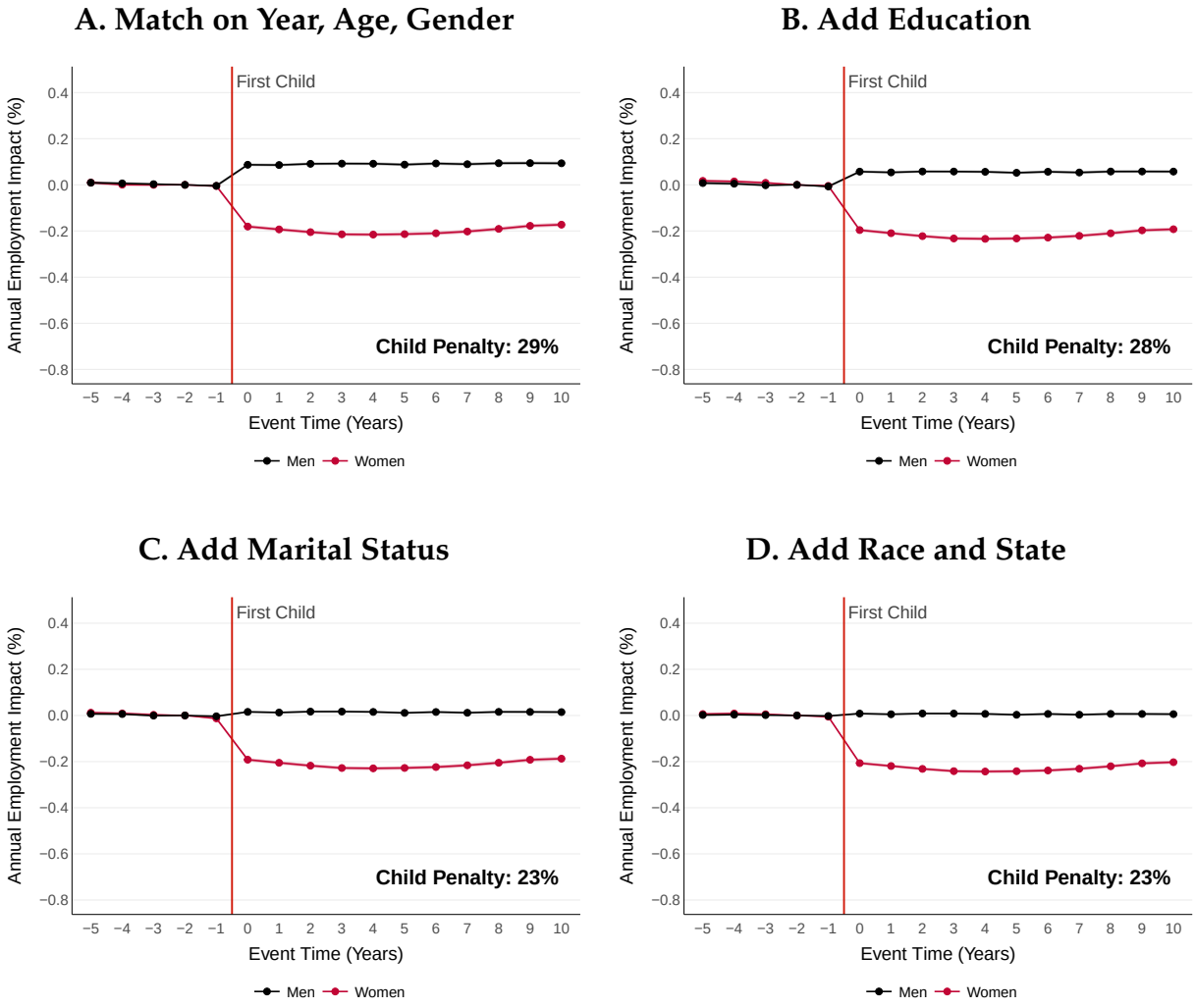
**TABLE A.1: TEST OF ASSUMPTION 3 IN PANEL DATA**
BALANCE BETWEEN ACTUAL AND SYNTHETIC OBSERVATIONS OF NEGATIVE EVENT TIMES

| | **Panel:** Observed Values | **Pseudo-Panel:** Synthetic Values | **Difference** |
|---|---|---|---|
| **Panel A: Men** | | | |
| *Annual Employment Rate* | | | |
| Event Time $\tau$ = -2 | 0.929 (0.007) | 0.917 (0.003) | -0.013 (0.008) |
| Event Times $\tau < 0$ | 0.927 (0.004) | 0.914 (0.002) | -0.013 (0.004) |
| *Weekly Employment Rate* | | | |
| Event Time $\tau$ = -2 | 0.923 (0.007) | 0.915 (0.003) | -0.009 (0.008) |
| Event Times $\tau < 0$ | 0.909 (0.004) | 0.914 (0.002) | 0.005 (0.005) |
| *Earnings (Normalized)* | | | |
| Event Time $\tau$ = -2 | 0.681 (0.019) | 0.670 (0.010) | -0.011 (0.021) |
| Event Times $\tau < 0$ | 0.675 (0.012) | 0.678 (0.009) | 0.003 (0.016) |
| **Panel B: Women** | | | |
| *Annual Employment Rate* | | | |
| Event Time $\tau$ = -2 | 0.939 (0.008) | 0.940 (0.004) | 0.000 (0.010) |
| Event Times $\tau < 0$ | 0.949 (0.004) | 0.942 (0.003) | -0.007 (0.005) |
| *Weekly Employment Rate* | | | |
| Event Time $\tau$ = -2 | 0.898 (0.010) | 0.881 (0.005) | -0.017 (0.011) |
| Event Times $\tau < 0$ | 0.896 (0.005) | 0.882 (0.003) | -0.014 (0.006) |
| *Earnings (Normalized)* | | | |
| Event Time $\tau$ = -2 | 0.728 (0.023) | 0.699 (0.011) | -0.029 (0.024) |
| Event Times $\tau < 0$ | 0.742 (0.013) | 0.704 (0.009) | -0.038 (0.015) |
| **Panel C: Men Minus Women** | | | |
| *Annual Employment Rate* | | | |
| Event Time $\tau$ = -2 | -0.010 (0.010) | -0.023 (0.005) | -0.013 (0.012) |
| Event Times $\tau < 0$ | -0.022 (0.005) | -0.028 (0.003) | -0.006 (0.006) |
| *Weekly Employment Rate* | | | |
| Event Time $\tau$ = -2 | 0.025 (0.013) | 0.034 (0.005) | 0.008 (0.014) |
| Event Times $\tau < 0$ | 0.013 (0.006) | 0.032 (0.003) | 0.019 (0.007) |
| *Earnings (Normalized)* | | | |
| Event Time $\tau$ = -2 | -0.047 (0.028) | -0.029 (0.015) | 0.018 (0.032) |
| Event Times $\tau < 0$ | -0.067 (0.018) | -0.026 (0.012) | 0.041 (0.021) |

Notes: This table tests Assumption 3 using pooled PSID/NLSY data. This assumption requires balance between actual and synthetic observations at negative event times, where the synthetic observations are obtained from the matching algorithm described in section 3.1. The table reports mean outcomes in the panel and pseudo-panel, as well as the differences between the two. Outcomes are shown for men (Panel A), women (Panel B), and men relative to women (Panel C). Each panel includes three outcomes: annual employment, weekly employment, and earnings (normalized by average earnings across all years, for men and women separately). Results are shown for event time $\tau = -2$ (omitted base year) and for the average of all negative event times. The findings support the empirical approach: the differences between the panel and pseudo-panel are small and mostly statistically insignificant. Standard errors are based on bootstrapping with replacement.

## FIGURE A.1: PSEUDO-EVENT STUDIES WITH DIFFERENT MATCHING VARIABLES

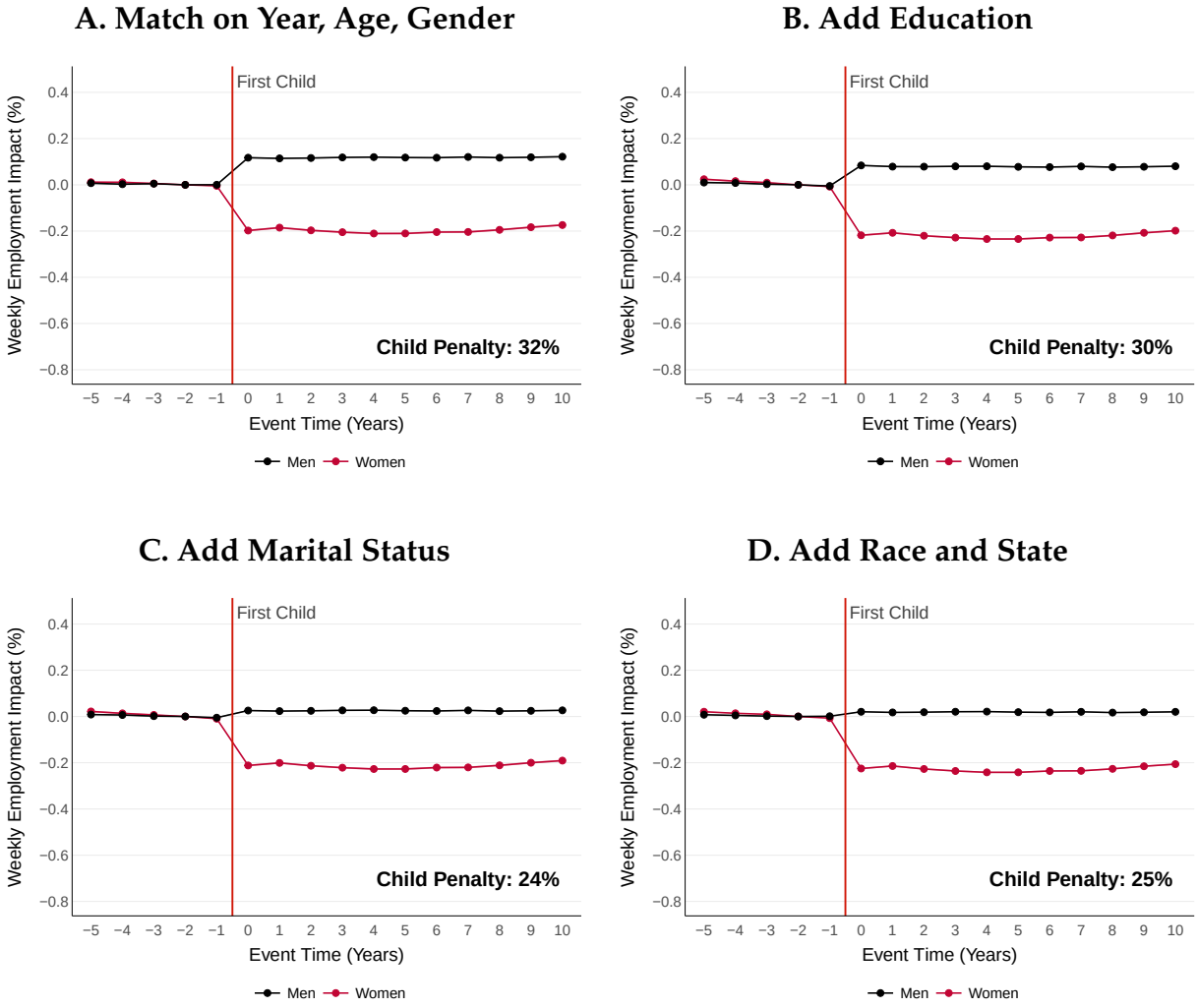
ANNUAL EMPLOYMENT




Notes: This figure presents pseudo-event studies of first childbirth for annual employment based on increasingly granular matching specifications. Panel A matches only on year, age, and gender, Panel B adds education, Panel C adds marital status, and Panel D adds race and state of residence. Panel D corresponds to the baseline specification presented in Figure 1. The more parsimonious specifications in Panels A-C are associated with selection bias, evidenced by the positive jumps for men between event times $\tau = -1$ and $\tau = 0$ as well as the discrepancy between these specifications and the true event study in Figure 1. The baseline specification in Panel D eliminates these selection problems.

## FIGURE A.2: PSEUDO-EVENT STUDIES WITH DIFFERENT MATCHING VARIABLES

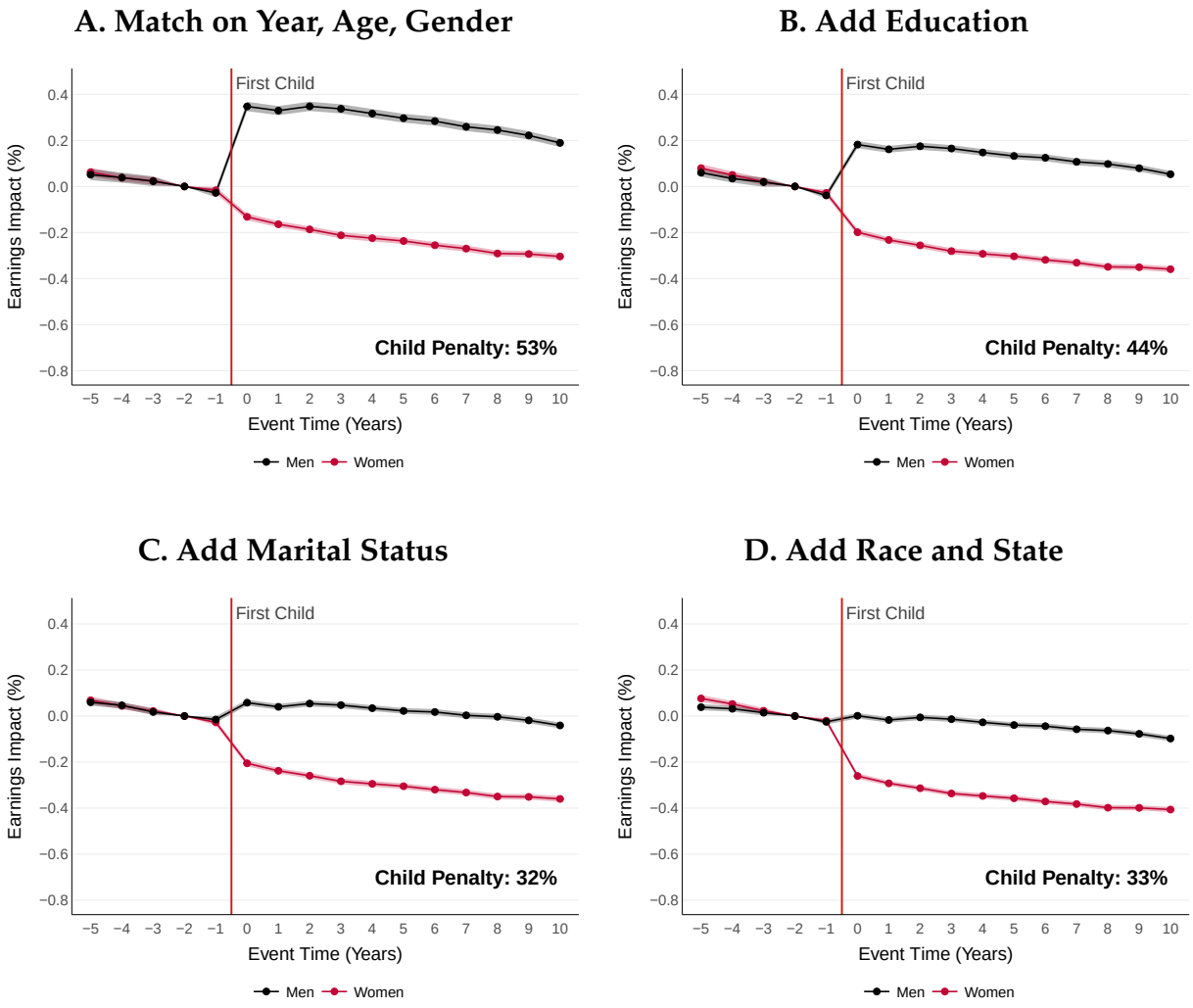
WEEKLY EMPLOYMENT




Notes: This figure presents pseudo-event studies of first childbirth for weekly employment based on increasingly granular matching specifications. Panel A matches only on year, age, and gender, Panel B adds education, Panel C adds marital status, and Panel D adds race and state of residence. Panel D corresponds to the baseline specification presented in Figure 1. The more parsimonious specifications in Panels A-C are associated with selection bias, evidenced by the positive jumps for men between event times $\tau = -1$ and $\tau = 0$ as well as the discrepancy between these specifications and the true event study in Figure 1. The baseline specification in Panel D eliminates these selection problems.

## FIGURE A.3: PSEUDO-EVENT STUDIES WITH DIFFERENT MATCHING VARIABLES

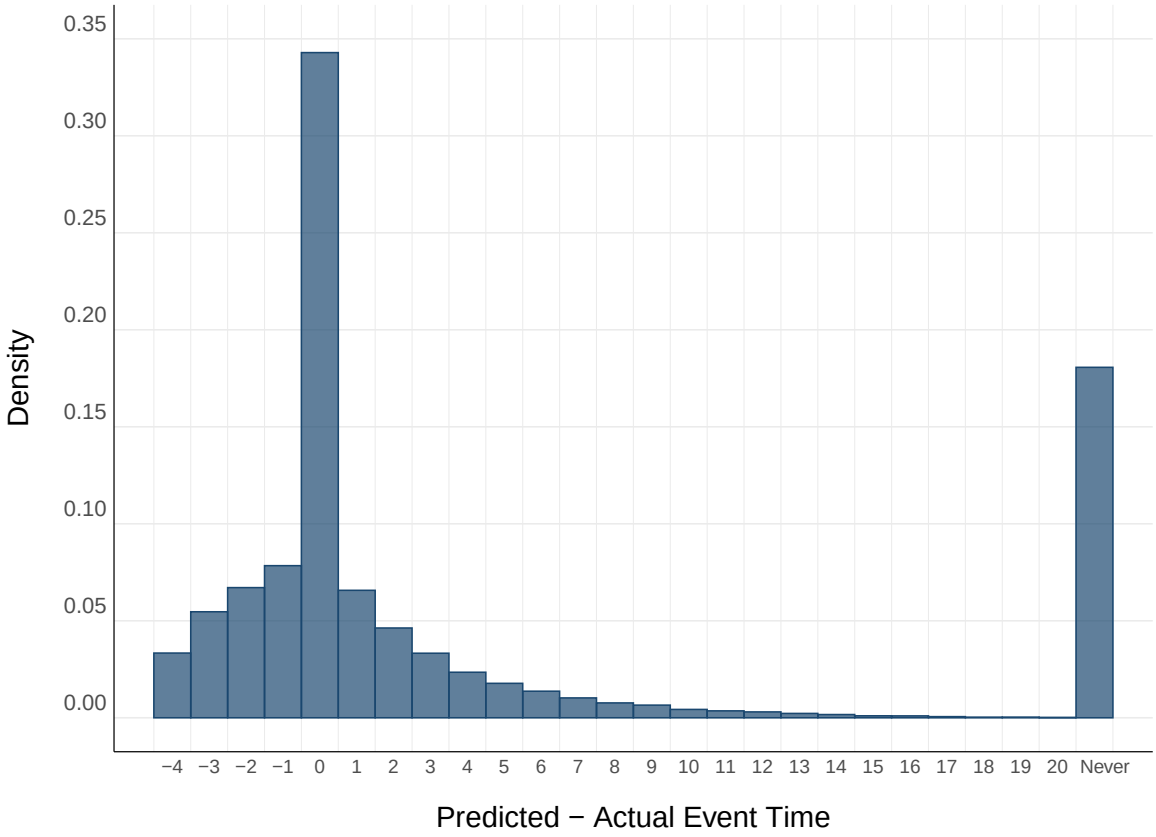
EARNINGS




Notes: This figure presents pseudo-event studies of first childbirth for earnings based on increasingly granular matching specifications. Panel A matches only on year, age, and gender, Panel B adds education, Panel C adds marital status, and Panel D adds race and state of residence. Panel D corresponds to the baseline specification presented in Figure 1. The more parsimonious specifications in Panels A-C are associated with selection bias, evidenced by the positive jumps for men between event times $\tau = -1$ and $\tau = 0$ as well as the discrepancy between these specifications and the true event study in Figure 1. The baseline specification in Panel D eliminates these selection problems.

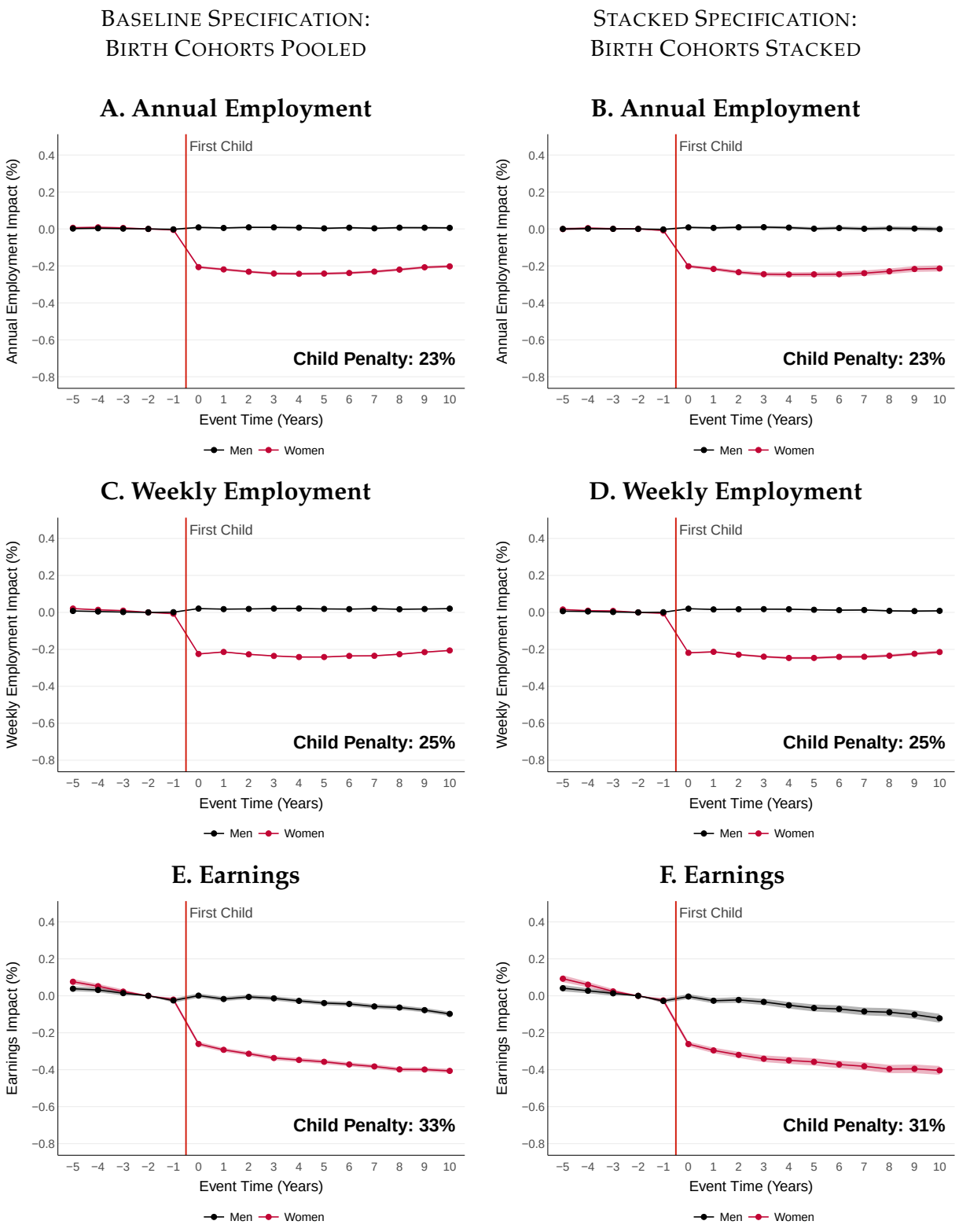
**FIGURE A.4: QUALITY OF FERTILITY PREDICTION IN PSEUDO-EVENT STUDY APPROACH**

PREDICTED VS ACTUAL EVENT TIMES AMONG CHILDLESS PEOPLE




Notes: This figure shows the distribution of within-person differences in predicted and actual event times among those observed without children. The distribution is based on panel data from PSID and NLSY between 1968-2019, sampling individuals observed after age 45 for whom completed fertility can be measured. Predicted event times for childless individuals are based on the matching specification used in the pseudo-event study approach (these event times vary from -5 to -1), while the actual event times for the same individuals are directly observed in the panel data. Event time is perfectly predicted for 34% of the data and with an error of less than four years for 74% of the data. The bin labeled "never" includes matched individuals (assigned to event times between -5 and -1) who never have children.

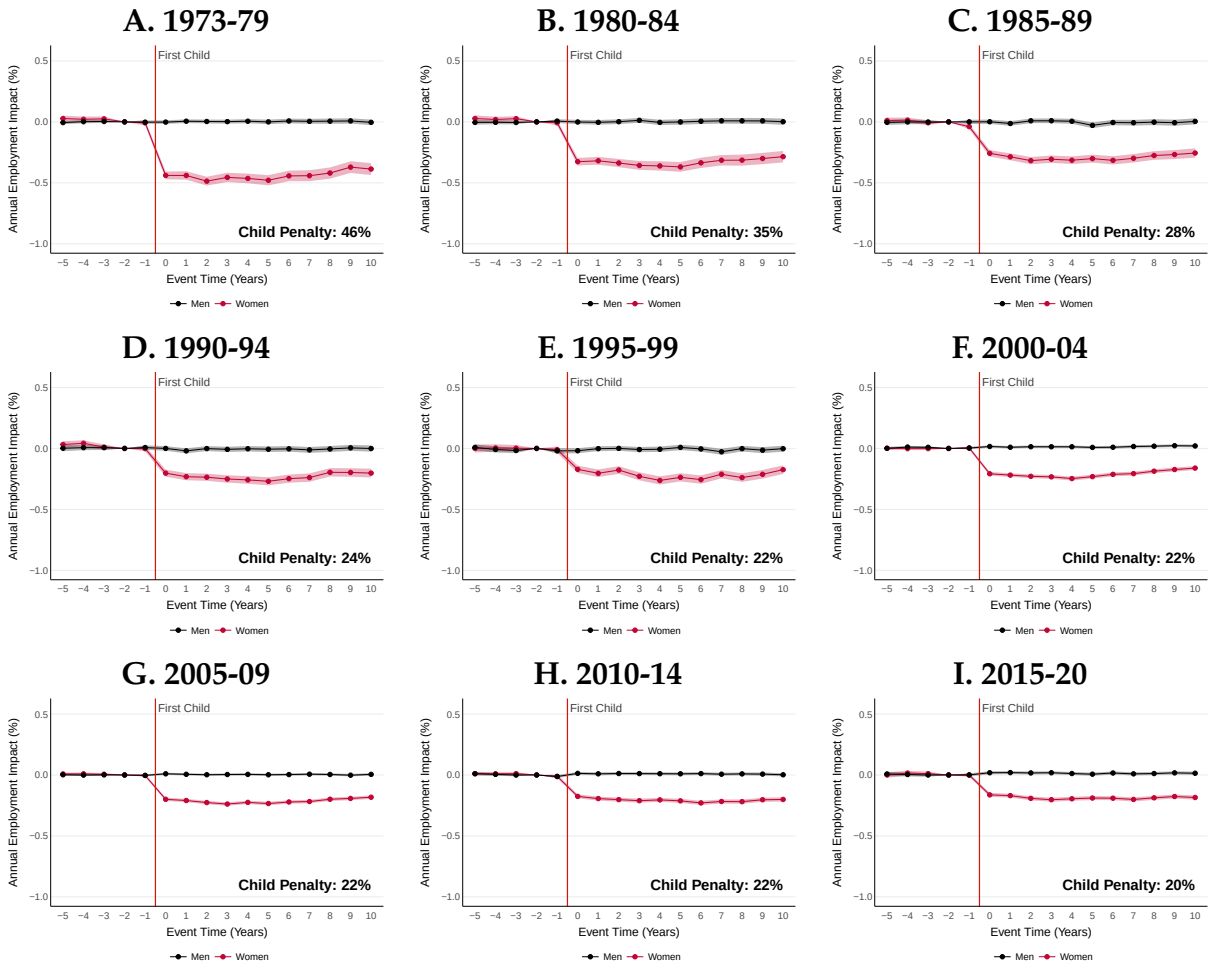
## FIGURE A.5: ROBUSTNESS TO HETEROGENEOUS TREATMENT EFFECTS




Notes: This figure investigates the possibility of bias from treatment-effect heterogeneity by comparing results from the baseline event study specification (pooling all birth cohorts) to results from a stacked event study specification (stacking different birth cohorts). Specifically, the stacked specification allows for heterogeneous effects by age at first birth (as specified in equation 9) and calculates a weighted average treatment effect using the sample shares of each cohort (as specified in equation 10). The baseline and stacked specifications produce almost identical results in all three labor market outcomes, indicating that heterogeneous treatment effects do not create bias in this context.

## FIGURE A.6: EVENT STUDIES OF FIRST CHILDBIRTH OVER TIME

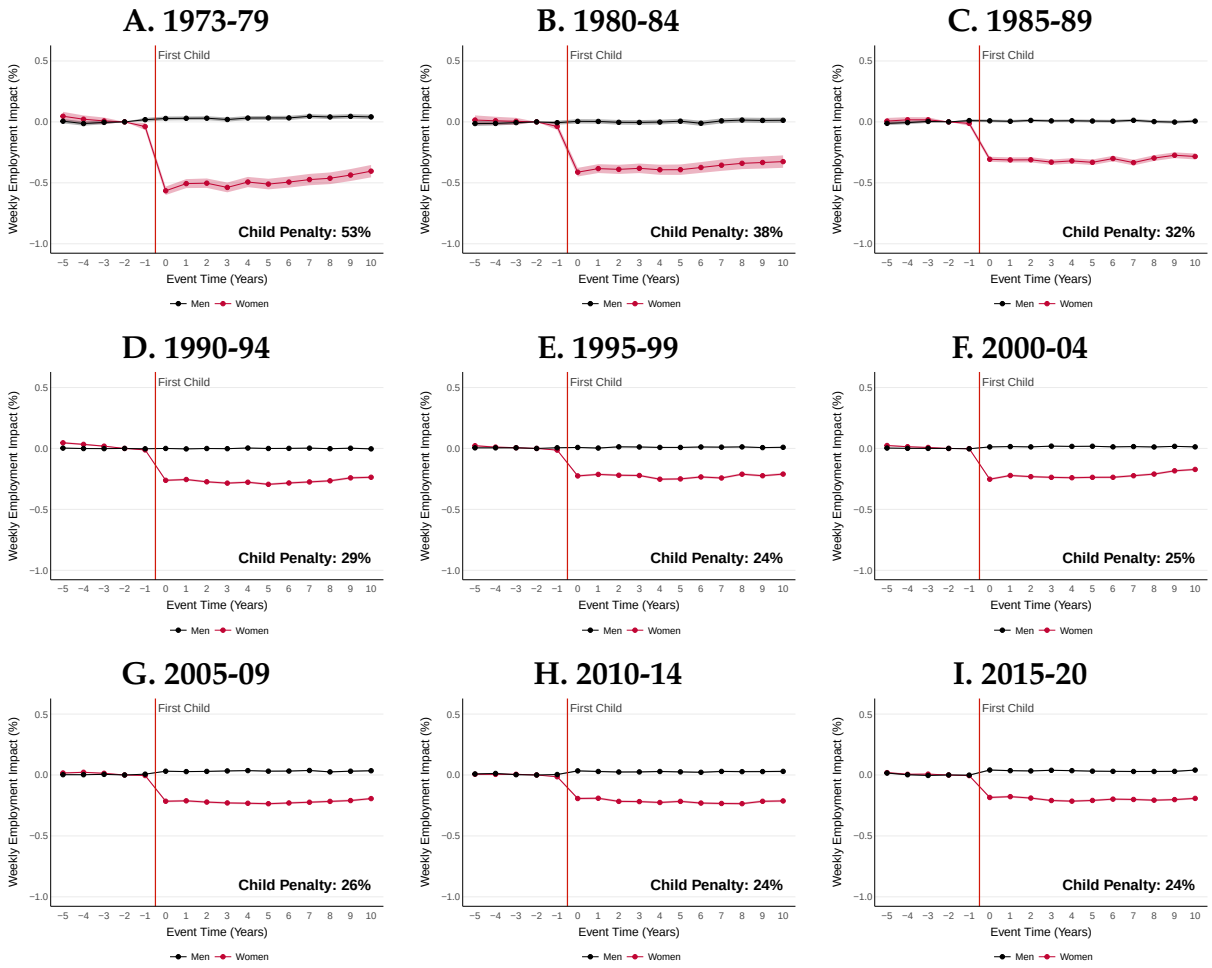
ANNUAL EMPLOYMENT




Notes: This figure shows event studies of first childbirth for annual employment in different time periods. The sample of parents is split by interview year and the event study specification (6) is run separately for each time period. The event studies start in 1973, because the first five years of the data (1968-1972) are reserved for obtaining synthetic pre-birth observations for those who had their first child in 1973. Each panel displays the average child penalty over event times 0-10 (defined in equation 8) for the time period in question. The 95% confidence intervals are based on robust standard errors.

## FIGURE A.7: EVENT STUDIES OF FIRST CHILDBIRTH OVER TIME

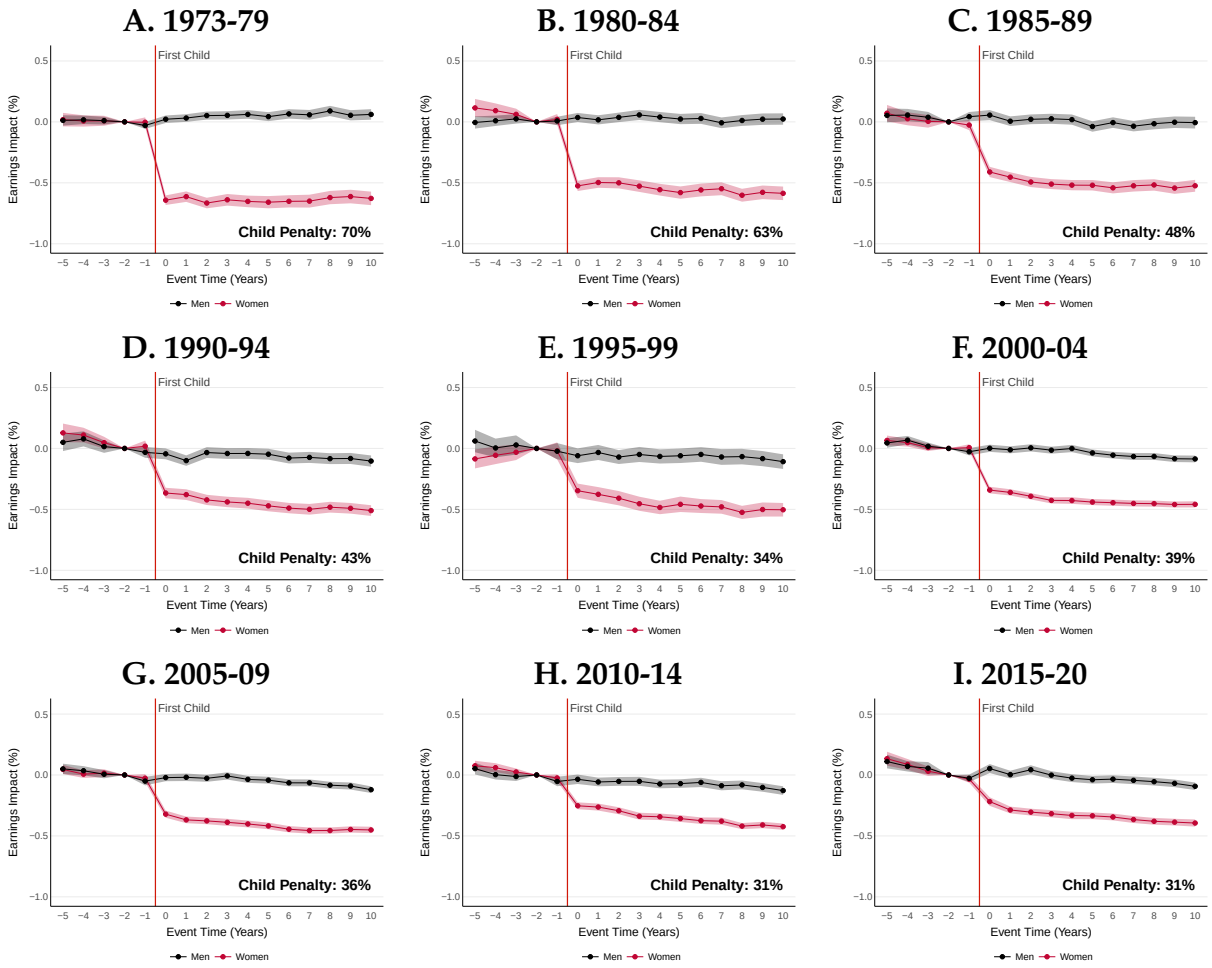
WEEKLY EMPLOYMENT




Notes: This figure shows event studies of first childbirth for weekly employment in different time periods. The sample of parents is split by interview year and the event study specification (6) is run separately for each time period. The event studies start in 1973, because the first five years of the data (1968-1972) are reserved for obtaining synthetic pre-birth observations for those who had their first child in 1973. Each panel displays the average child penalty over event times 0-10 (defined in equation 8) for the time period in question. The 95% confidence intervals are based on robust standard errors.

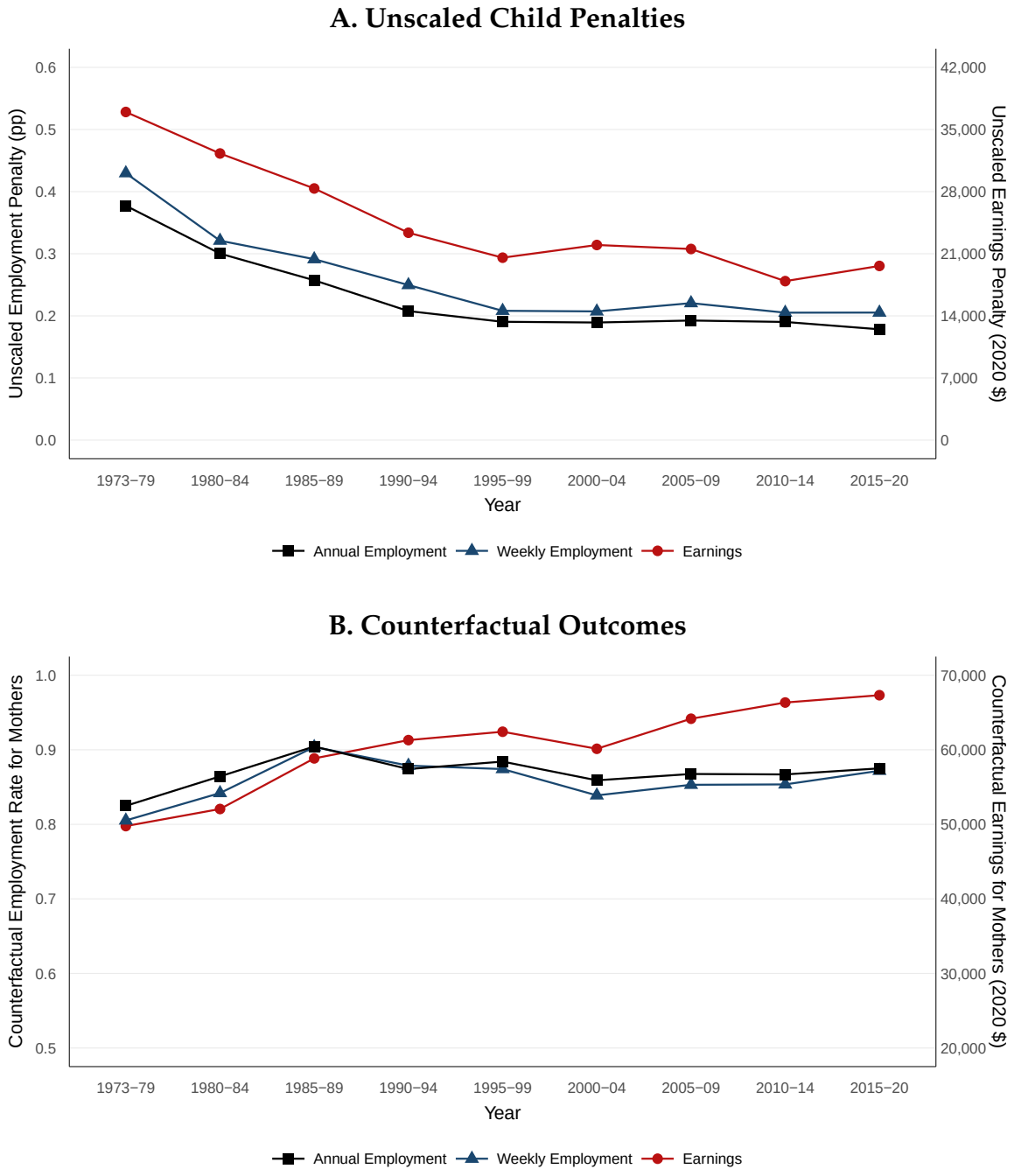
## FIGURE A.8: EVENT STUDIES OF FIRST CHILDBIRTH OVER TIME

EARNINGS



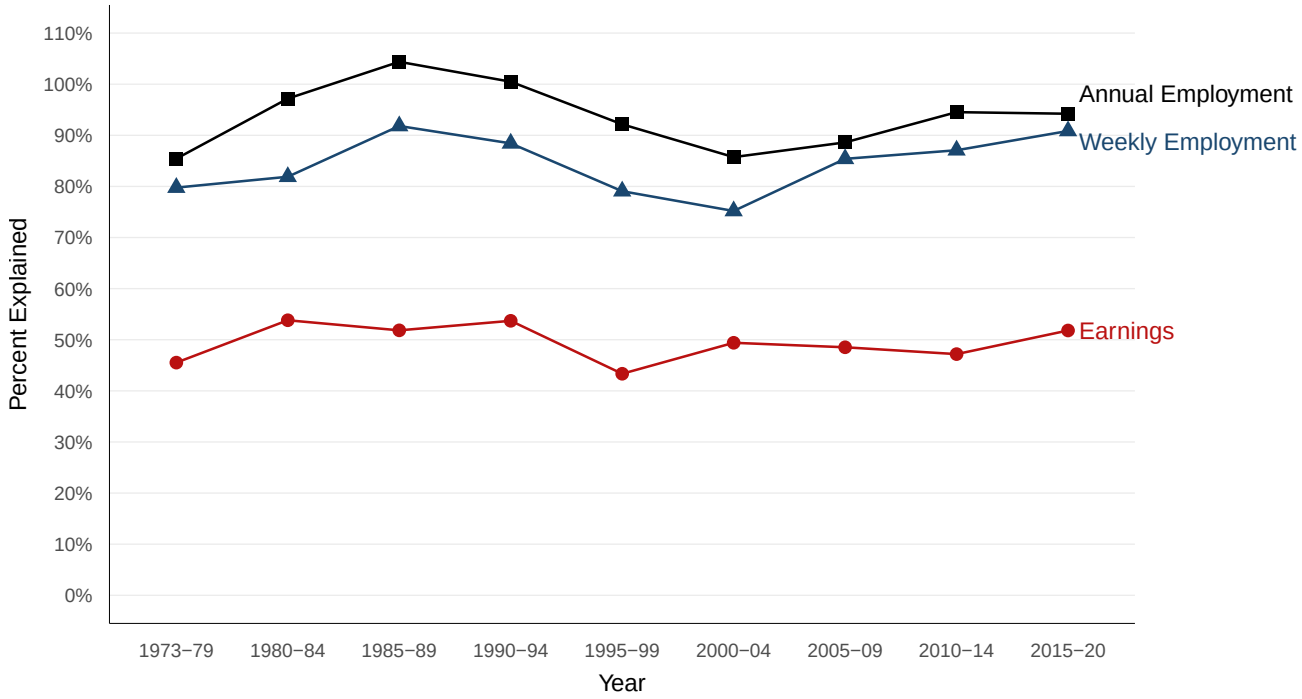

Notes: This figure shows event studies of first childbirth for earnings in different time periods. The sample of parents is split by interview year and the event study specification (6) is run separately for each time period. The event studies start in 1973, because the first five years of the data (1968-1972) are reserved for obtaining synthetic pre-birth observations for those who had their first child in 1973. Each panel displays the average child penalty over event times 0-10 (defined in equation 8) for the time period in question. The 95% confidence intervals are based on robust standard errors.

## FIGURE A.9: UNSCALED CHILD PENALTIES AND COUNTERFACTUALS OVER TIME




Notes: This figure investigates if the time series of scaled child penalties (effects in percentage terms) shown in Figure 4 are driven primarily by changes in unscaled child penalties (effects in absolute terms) or by changes in the scaling factor (the level of the counterfactual outcome). Panel A shows the evolution of unscaled child penalties (employment effects in percentage points and earnings effects in dollars), while Panel B hows the evolution of the counterfactual levels used for scaling. The earnings estimates have been adjusted to 2020 dollars using nominal earnings growth in the full sample of working-age men and women in CPS data. The figure shows that the evolution of unscaled penalties is similar to the evolution of scaled penalties: a decline until the mid-1990s and then stagnation. The counterfactual outcomes have remained relatively constant for employment, while they have increased gradually for earnings. Hence, changes in the baseline hardly matter for the evolution of scaled employment penalties, while they play some role for the evolution of scaled earnings penalties.

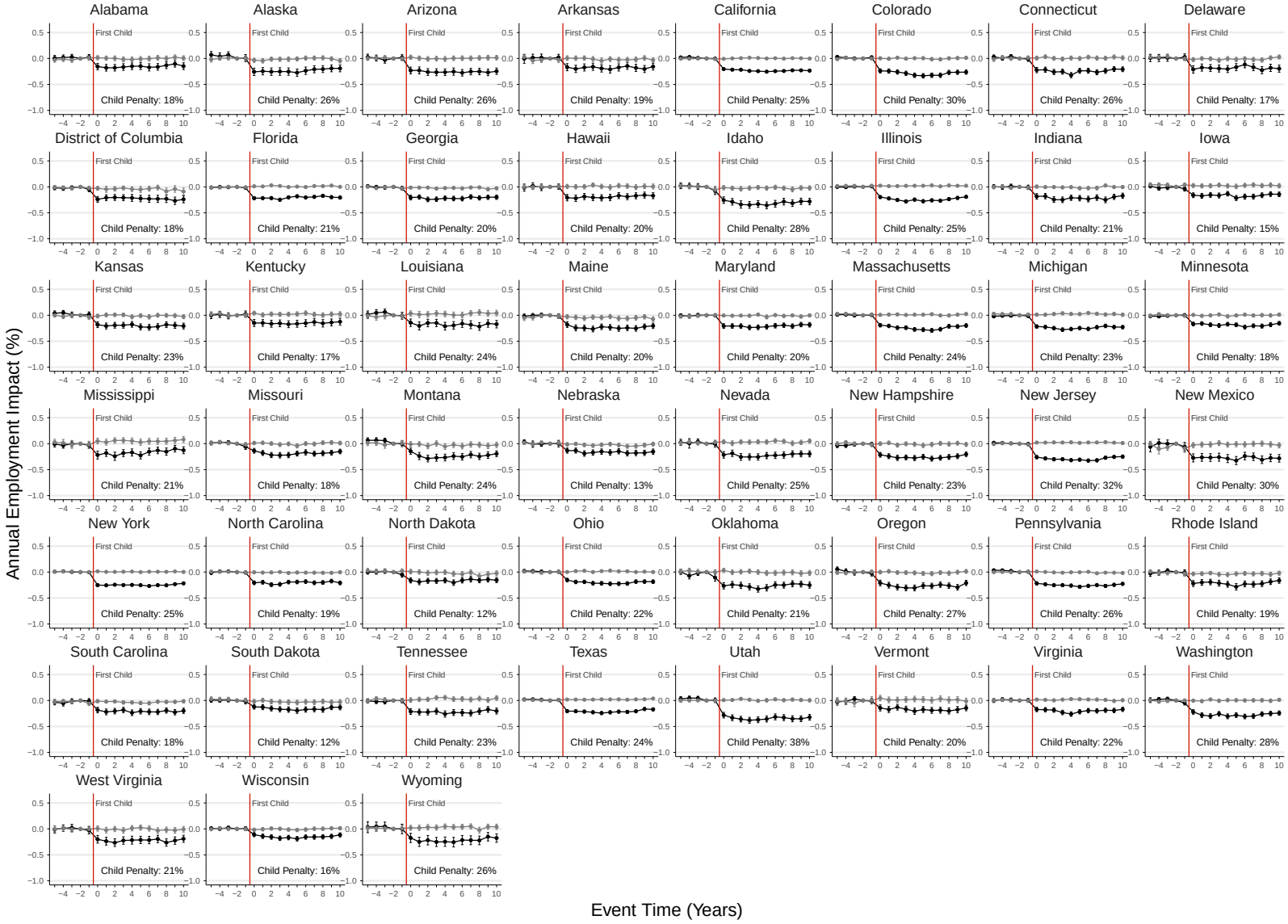
## FIGURE A.10: FRACTION OF RAW GENDER GAP EXPLAINED BY CHILD PENALTIES




Notes: This figure shows the fraction of the raw gender gap for parents explained by child penalties over time. Results are shown for each of the three labor market outcomes: annual employment, weekly employment, and earnings. The raw gender gap is defined as the percentage difference between men and women with children, and the child penalty estimates are shown in Figure 4.

## FIGURE A.11: EVENT STUDIES OF FIRST CHILDBIRTH ACROSS STATES

### ANNUAL EMPLOYMENT




Notes: This figure shows event studies of first childbirth in annual employment for each of the 51 US states (including the federal district of D.C.). State-level event studies are constructed by interacting the event time dummies in equation (6) with state dummies, estimating percentage impacts of childbirth on men and women at each event time ($\hat{P}_\tau^m$ and $\hat{P}_\tau^w$) as well as average child penalties over event times 0-10 separately for each state. Men are shown in gray and women are shown in black. In this specification, the lifecycle and time trends in equation (6) are estimated at the level of census divisions. The 95% confidence intervals are based on robust standard errors.

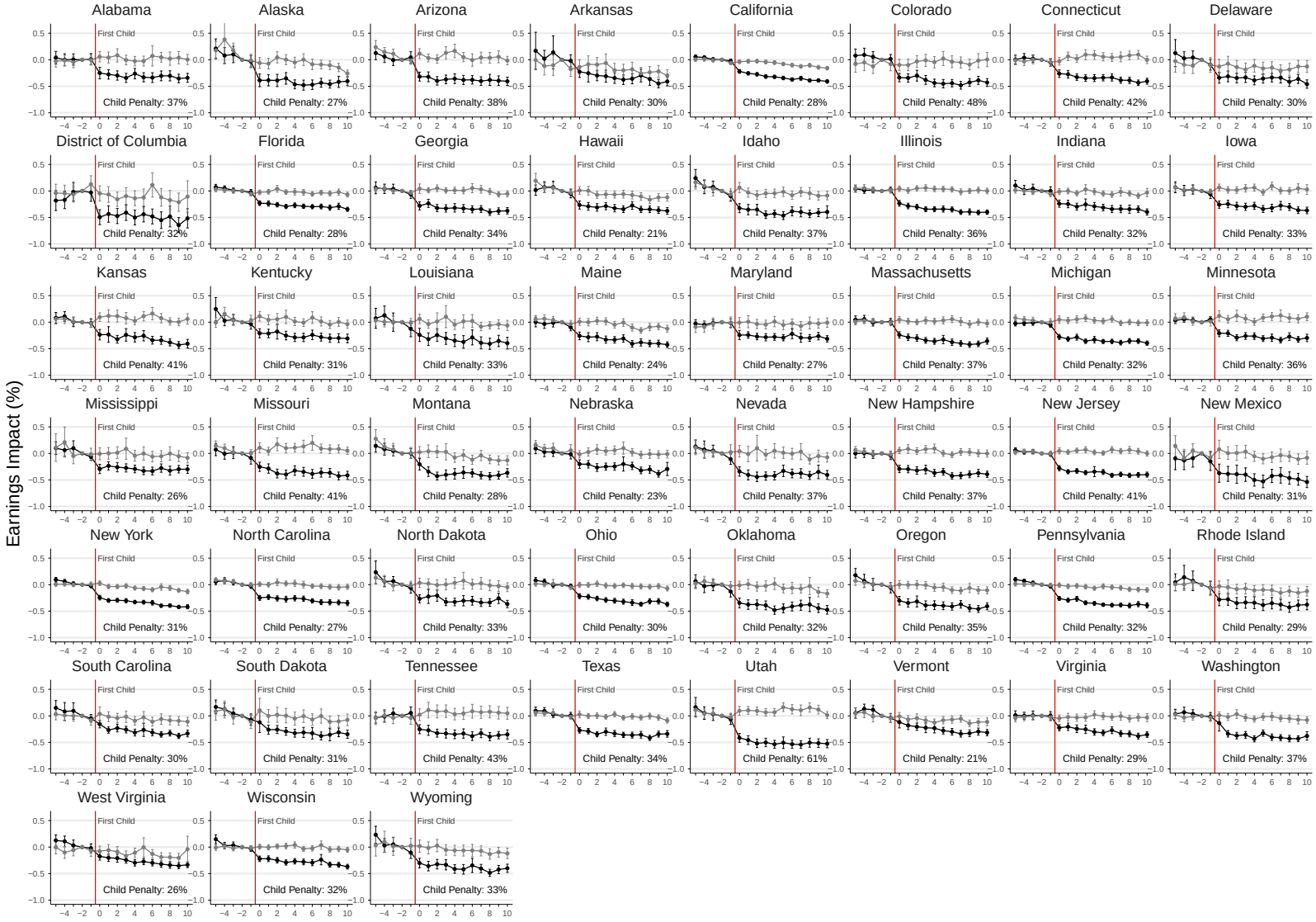
## FIGURE A.12: EVENT STUDIES OF FIRST CHILDBIRTH ACROSS STATES

WEEKLY EMPLOYMENT


Notes: This figure shows event studies of first childbirth in weekly employment for each of the 51 US states (including the federal district of D.C.). State-level event studies are constructed by interacting the event time dummies in equation (6) with state dummies, estimating percentage impacts of childbirth on men and women at each event time ($\hat{P}^m_\tau$ and $\hat{P}^w_\tau$) as well as average child penalties over event times 0-10 separately for each state. Men are shown in gray and women are shown in black. In this specification, the lifecycle and time trends in equation (6) are estimated at the level of census divisions. The 95% confidence intervals are based on robust standard errors.

## FIGURE A.13: EVENT STUDIES OF FIRST CHILDBIRTH ACROSS STATES

EARNINGS




Notes: This figure shows event studies of first childbirth in earnings for each of the 51 US states (including the federal district of D.C.). State-level event studies are constructed by interacting the event time dummies in equation (6) with state dummies, estimating percentage impacts of childbirth on men and women at each event time ($\hat{P}_\tau^m$ and $\hat{P}_\tau^w$) as well as average child penalties over event times 0-10 separately for each state. Men are shown in gray and women are shown in black. In this specification, the lifecycle and time trends in equation (6) are estimated at the level of census divisions. The 95% confidence intervals are based on robust standard errors.

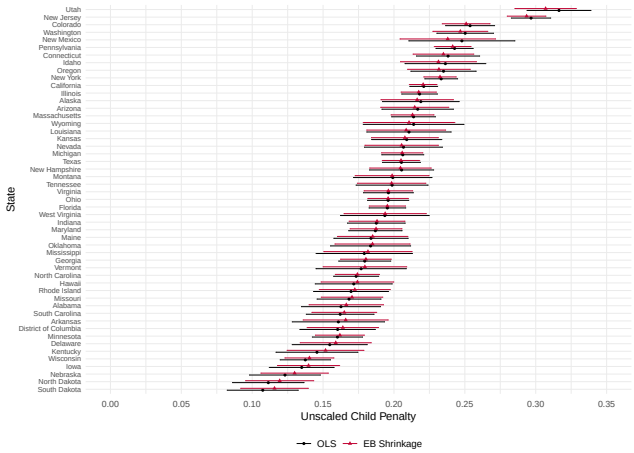
## FIGURE A.14: CHILD PENALTIES VS RAW GENDER GAPS ACROSS STATES


Notes: This figure provides scatter plots of child penalties against raw gender gaps for parents across states. Results are shown for each of the three labor market outcomes: annual employment, weekly employment, and earnings. The raw gender gap is defined as the percentage difference between men and women with children, and the child penalty estimates for each outcome and state are shown in Figures A.11-A.13.

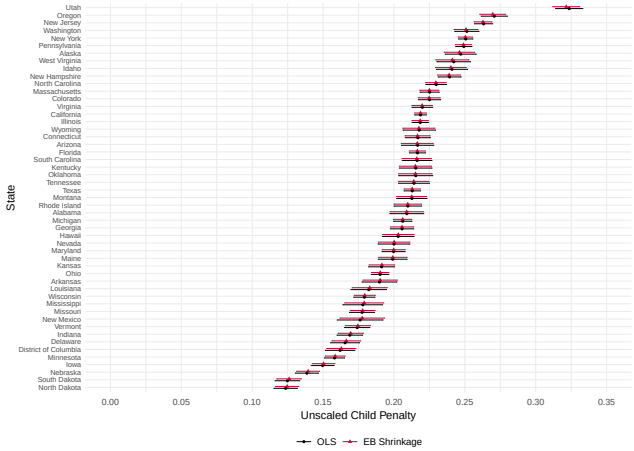
## FIGURE A.15: EB VS OLS ESTIMATES OF CHILD PENALTIES

POINT ESTIMATES AND CONFIDENCE INTERVALS BY STATE

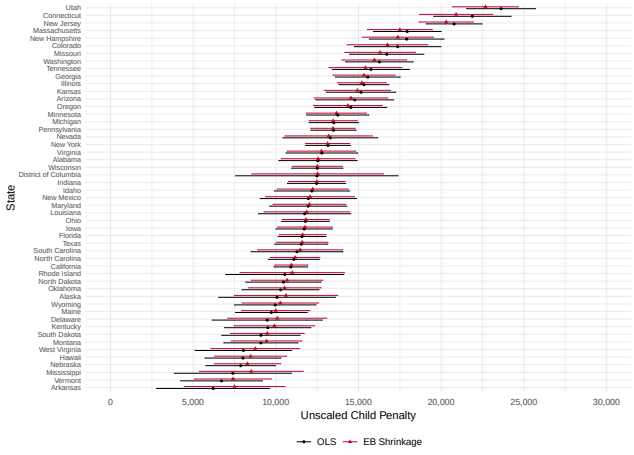
### A. Annual Employment




### B. Weekly Employment




### C. Earnings

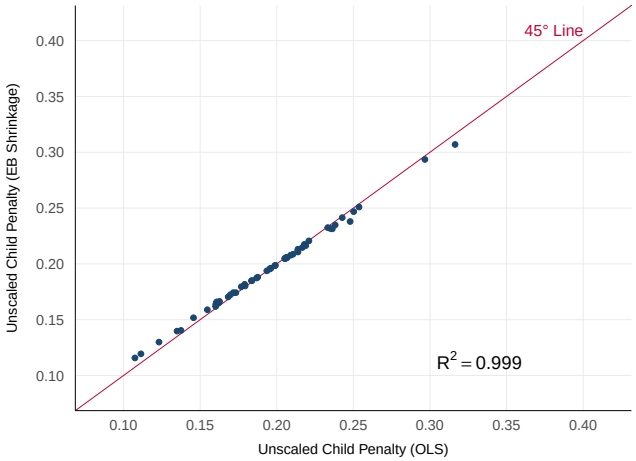



Notes: This figure compares OLS and Empirical Bayes (EB) estimates of unscaled child penalties for each state and each labor market outcome. The EB estimates are based on the linear-shrinkage formula in equation (13). The EB shrinkage adjustment hardly changes the estimates. The reason is the high statistical precision of the pseudo-event study approach: the imprecision of the state-level OLS estimates is very small compared to the variation in OLS estimates across states.

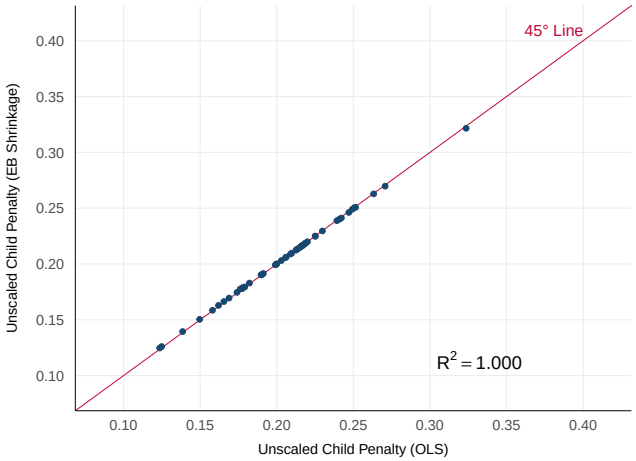
## FIGURE A.16: EB VS OLS ESTIMATES OF CHILD PENALTIES

SCATTERPLOTS

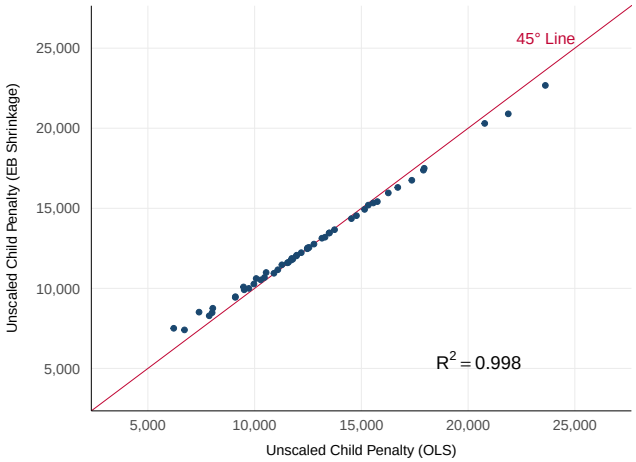
### A. Annual Employment




### B. Weekly Employment

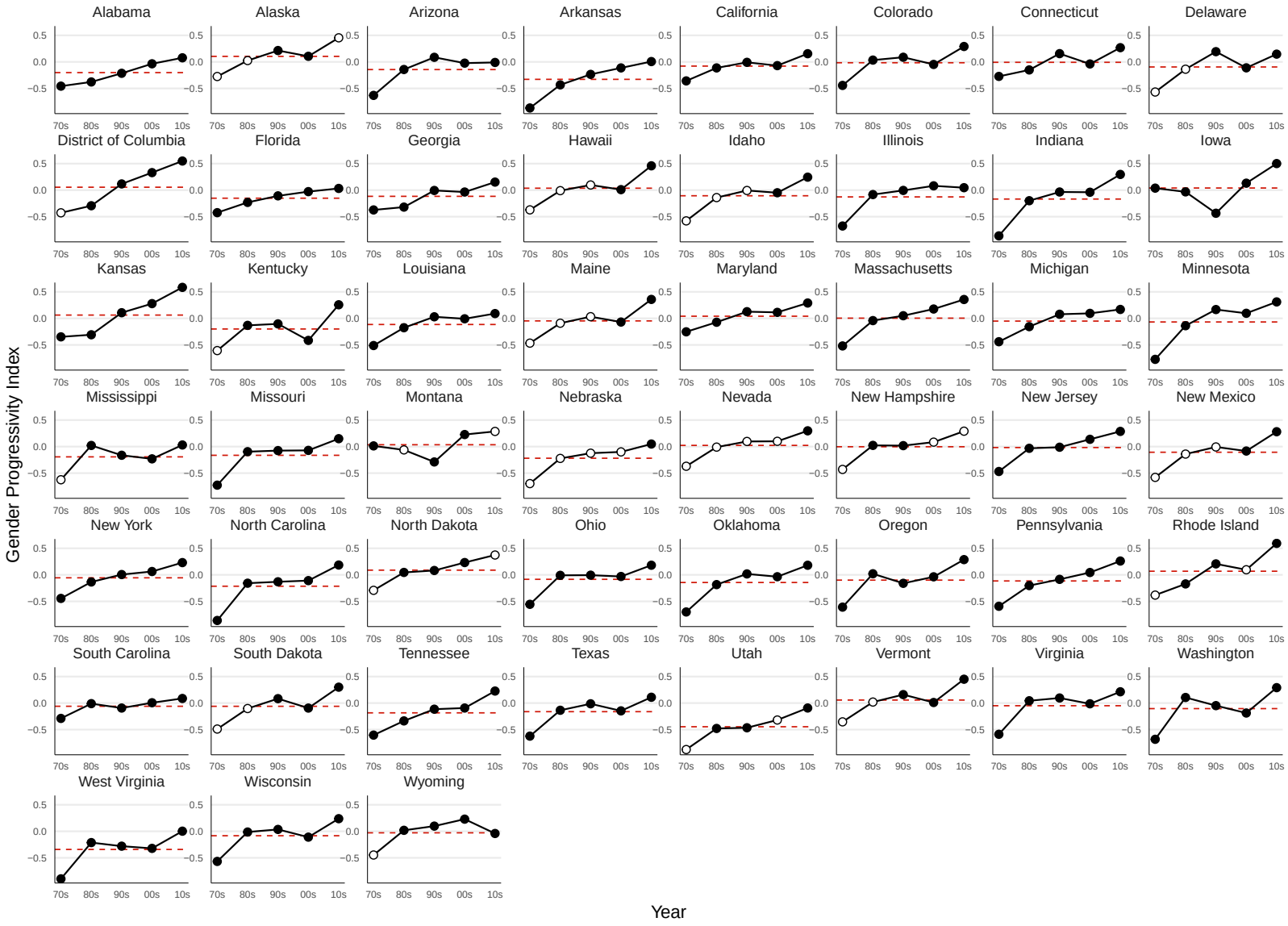



### C. Earnings




Notes: This figure plots Empirical Bayes (EB) estimates of child penalties against OLS estimates of child penalties across states for each labor market outcome. The EB estimates are based on the linear-shrinkage formula in equation (13). The EB-OLS pairs align almost perfectly with the 45-degree line for all three outcomes. The reason is the high statistical precision of the pseudo-event study approach: the imprecision of the state-level OLS estimates is very small compared to the variation in OLS estimates across states.

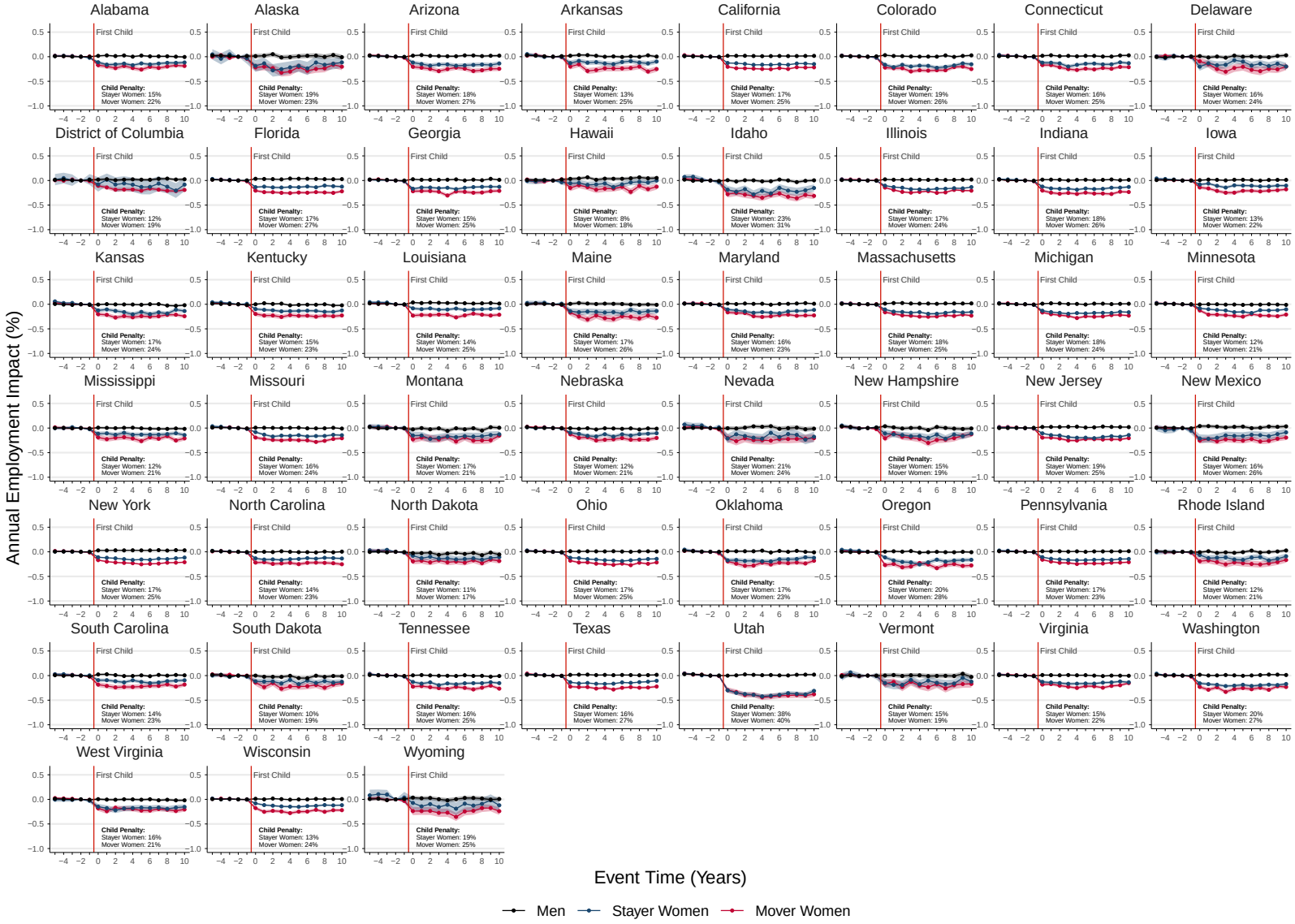
## FIGURE A.17: GENDER PROGRESSIVITY INDEX BY STATE AND TIME




Notes: This figure presents time series of the Gender Progressivity Index (GPI) in each state over the last five decades. Using GSS data from 1972-2018, the index is calculated as the average standardized response to questions that elicit attitudes towards gender roles in families with children. The standardization ensures that the index has mean zero and standard deviation one. Three gender norms questions available in all five decades of GSS data are included in the construction of the index. Because these questions were not asked in every state in every decade, some state-decade observations are missing. Missing state-decade observations have been imputed based on the percentile of the state's GPI in the decades where it is observed. Actual state-decade observations are indicated by filled dots and imputed observations are indicated by empty dots.

## FIGURE A.18: EVENT STUDIES OF FIRST CHILDBIRTH FOR MOVERS VS STAYERS BY STATE OF BIRTH

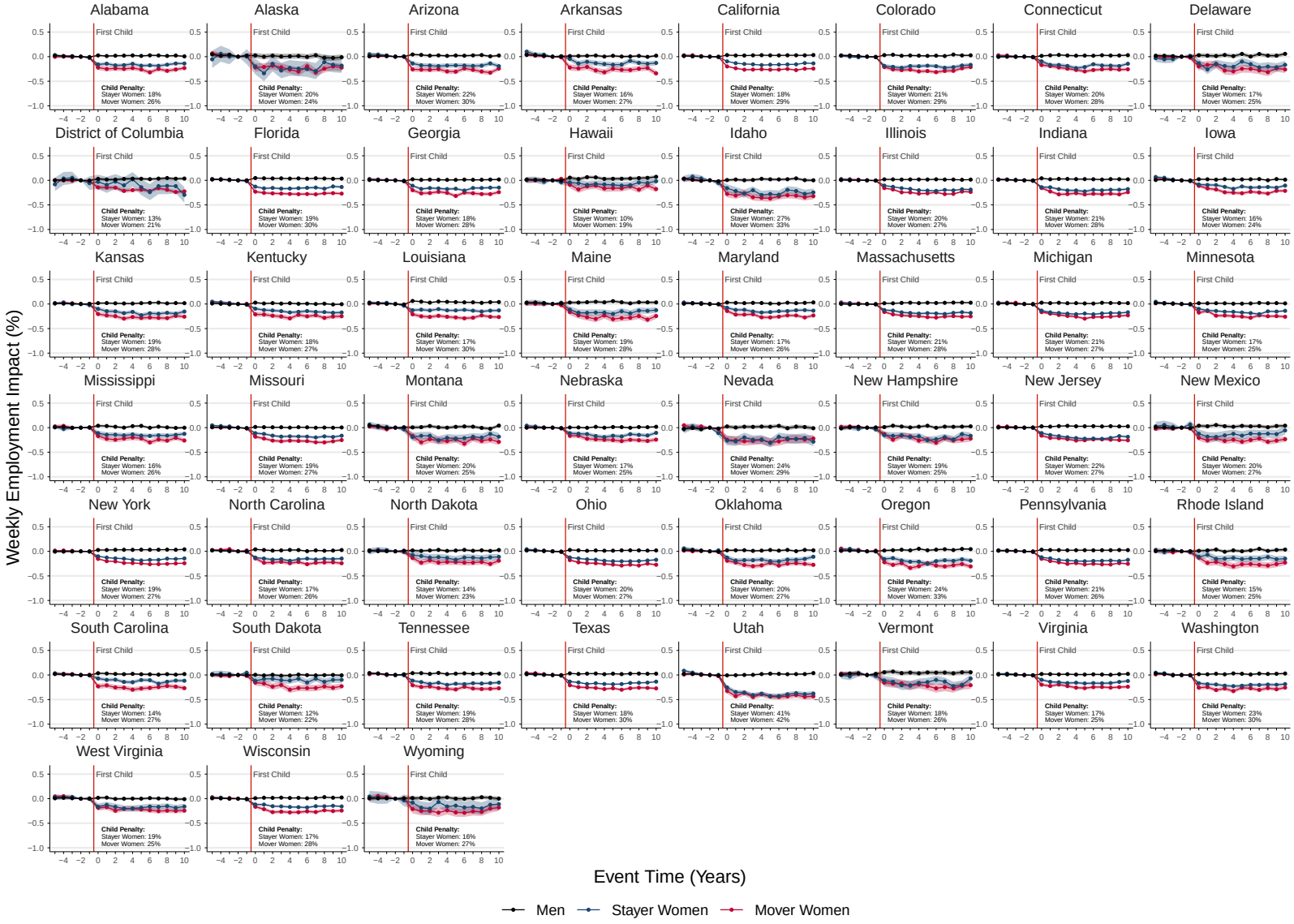
### ANNUAL EMPLOYMENT




Notes: This figure presents event studies of first childbirth for movers and stayers born in different states. Movers are defined as US-born individuals who reside in a different state than where they were born, while stayers are defined as US-born individuals who reside in the same state as where they were born. To construct the figure, specification (6) is run separately for women movers and women stayers, interacting the event time dummies by state-of-birth dummies. The sample of men is not split by mover/stayer status as childbirth is a non-event for them regardless of status. The outcome is annual employment. Each panel displays child penalties over event times 0-10 for mover women and stayer women with a given state of birth. The 95% confidence intervals are based on robust standard errors. The sample is based on ACS data from 2000-2019, which contains information on both state of residence and state of birth.

## FIGURE A.19: EVENT STUDIES OF FIRST CHILDBIRTH FOR MOVERS VS STAYERS BY STATE OF BIRTH

### WEEKLY EMPLOYMENT



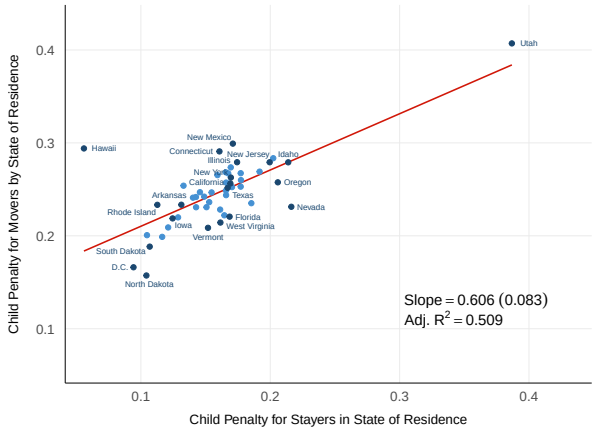

Notes: This figure presents event studies of first childbirth for movers and stayers born in different states. Movers are defined as US-born individuals who reside in a different state than where they were born, while stayers are defined as US-born individuals who reside in the same state as where they were born. To construct the figure, specification (6) is run separately for women movers and women stayers, interacting the event time dummies by state-of-birth dummies. The sample of men is not split by mover/stayer status as childbirth is a non-event for them regardless of status. The outcome is weekly employment. Each panel displays child penalties over event times 0-10 for mover women and stayer women with a given state of birth. The 95% confidence intervals are based on robust standard errors. The sample is based on ACS data from 2000-2019, which contains information on both state of residence and state of birth.

## FIGURE A.20: EPIDEMIOLOGICAL STUDY OF US MOVERS

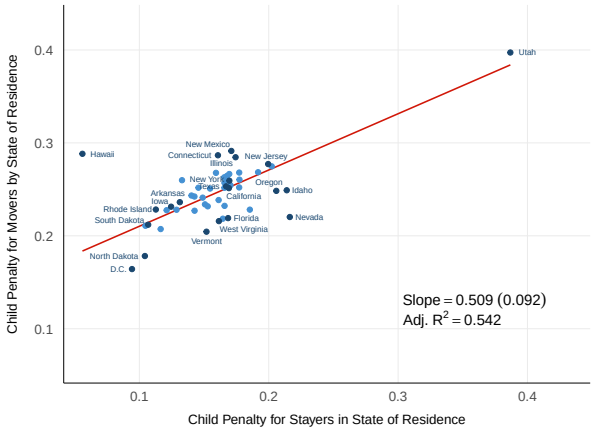
CHILD PENALTIES FOR MOVERS VS STAYERS BY STATE OF RESIDENCE

*Annual Employment*

**A. Raw Data**

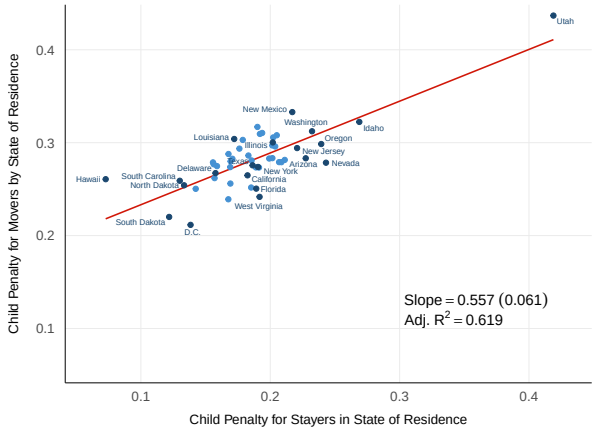



**B. Residualized by Effect of State-of-Birth Penalties**

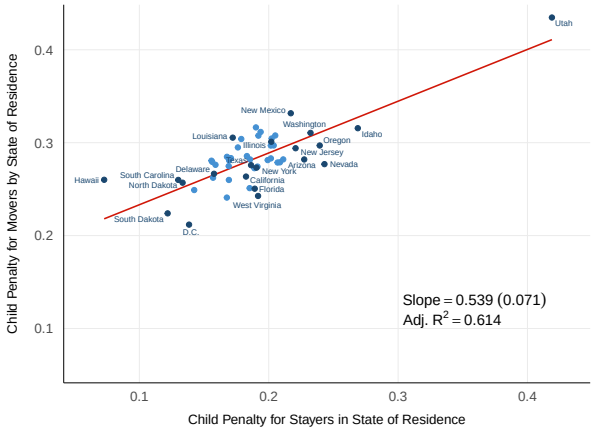



*Weekly Employment*

**C. Raw Data**



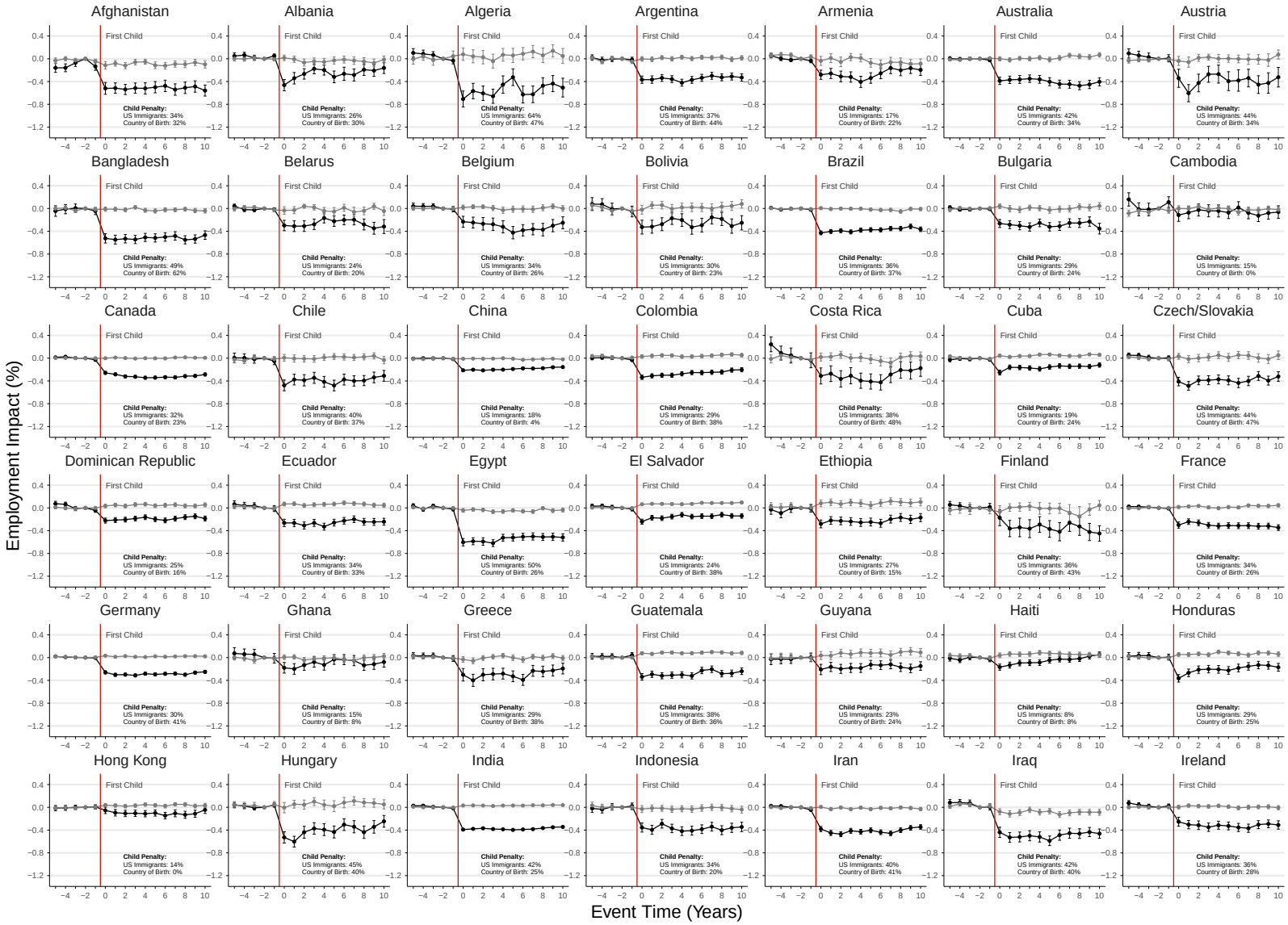

**D. Residualized by Effect of State-of-Birth Penalties**




Notes: This figure is symmetric to Figure 13, but focuses on the effect of residence state rather than the effect of birth state. Specifically, the figure provides scatter plots of the child penalty for movers against the child penalty for stayers by state of residence. Movers are defined as US-born individuals who reside in a different state than where they were born, while stayers are defined as US-born individuals who reside in the same state as where they were born. The left panels show raw child penalties, while the right panels show residualized child penalties using the specification in eq. (15). The residualized plots control for selection on state of birth, which would otherwise contaminate the estimated effects of state of residence (local labor markets) with effects of state of birth (norms/culture). The figure shows that place of residence has sizable effects, although not quite as strong as the effects of place of birth. The sample is based on ACS data from 2000-2019, which contains information on both state of residence and state of birth.

## FIGURE A.21: EPIDEMIOLOGICAL STUDY OF FOREIGN IMMIGRANTS

EVENT STUDIES OF FIRST CHILDBIRTH FOR IMMIGRANTS BY COUNTRY OF BIRTH




Notes: This figure presents event studies of first childbirth for foreign-born immigrants by country of birth. Each panel displays the child penalty for US immigrants (based on the series shown) and the child penalty in their country of birth (based on Kleven, Landais and Leite-Mariante 2025). The outcome is pooled employment (combining information on weekly and annual employment) and the sample is based on ACS data from 2000-2019 and CPS data from 1994-2020. Men are shown in gray and women are shown in black. The 95% confidence intervals are based on robust standard errors.

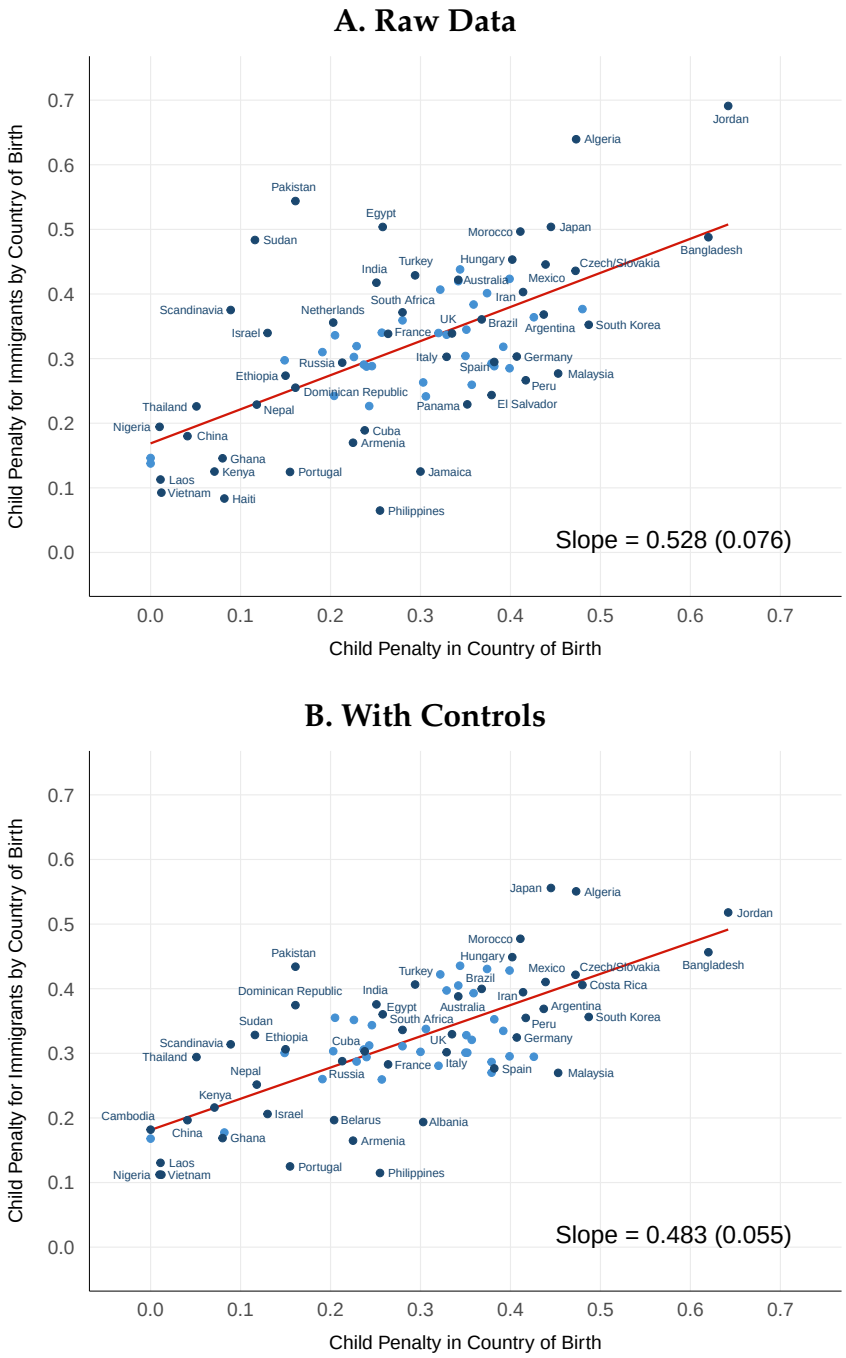
## FIGURE A.21: EPIDEMIOLOGICAL STUDY OF FOREIGN IMMIGRANTS (CONTINUED)

EVENT STUDIES OF FIRST CHILDBIRTH FOR IMMIGRANTS BY COUNTRY OF BIRTH


Notes: This figure presents event studies of first childbirth for foreign-born immigrants by country of birth. Each panel displays the child penalty for US immigrants (based on the series shown) and the child penalty in their country of birth (based on Kleven, Landais and Leite-Mariante 2025). The outcome is pooled employment (combining information on weekly and annual employment) and the sample is based on ACS data from 2000-2019 and CPS data from 1994-2020. Men are shown in gray and women are shown in black. The 95% confidence intervals are based on robust standard errors.

## FIGURE A.22: EPIDEMIOLOGICAL STUDY OF FOREIGN IMMIGRANTS

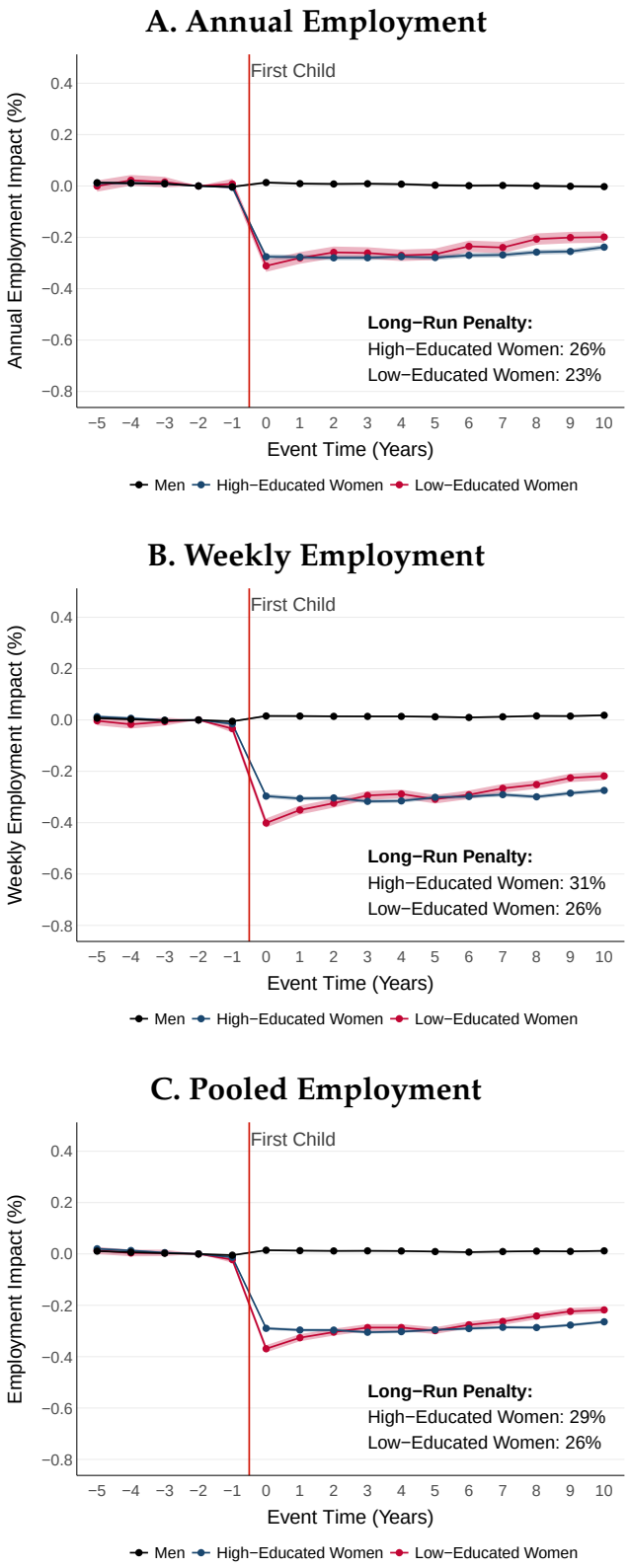
CHILD PENALTIES FOR IMMIGRANTS VS CHILD PENALTIES IN COUNTRIES OF BIRTH




Notes: This figure presents scatter plots of child penalties for foreign-born immigrants against child penalties in country of birth. The underlying event studies for US immigrants are shown in Appendix Figure A.21 and the child penalties in country of birth are taken from Kleven, Landais and Leite-Mariante (2025). Panel A shows raw child penalty estimates, while Panel B controls for differences in education, marriage, race, fertility, age at first birth, and US location across immigrants from different countries. The specification of these control variables corresponds to the variables shown in Table 3. To construct Panel B, immigrant penalties are regressed on birth-country penalties and demographic controls, residualizing the immigrant penalties by the estimated effect of the controls for each country. The average effect of controls across all countries is added back to the residualized outcome to make the levels in Panel A and B comparable. The outcome is pooled employment (combining information on weekly and annual employment) and the sample is based on ACS data from 2000-2019 and CPS data from 1994-2020.

## FIGURE A.23: EVENT STUDIES OF FIRST CHILDBIRTH BY EDUCATION

FOREIGN IMMIGRANTS

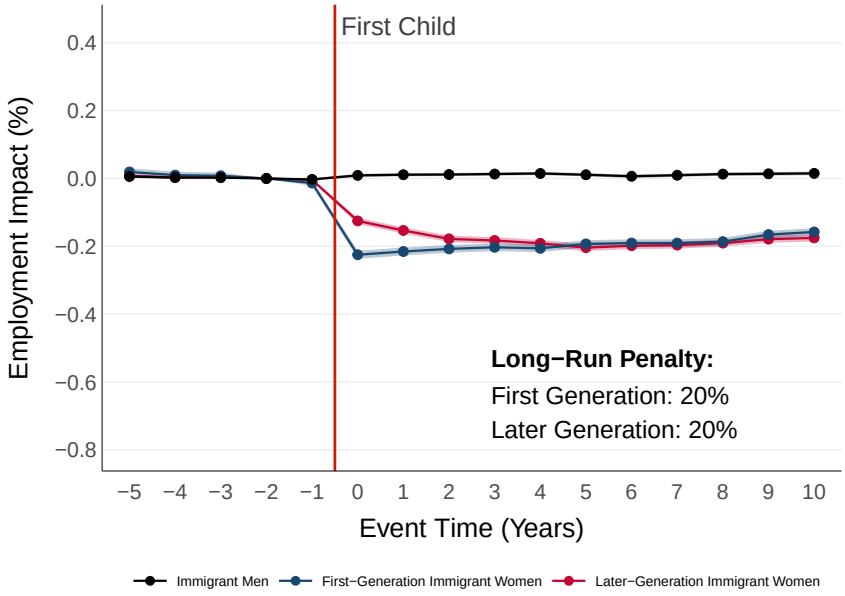



Notes: This figure presents event studies of first childbirth by female education level for foreign-born immigrants. The figure is constructed in the same way as the education part of Figure 8 for the full sample. Results are shown for three labor market outcomes: annual employment, weekly employment, and pooled employment. The analysis is based on ACS data from 2000-2019 and CPS data from 1994-2020.

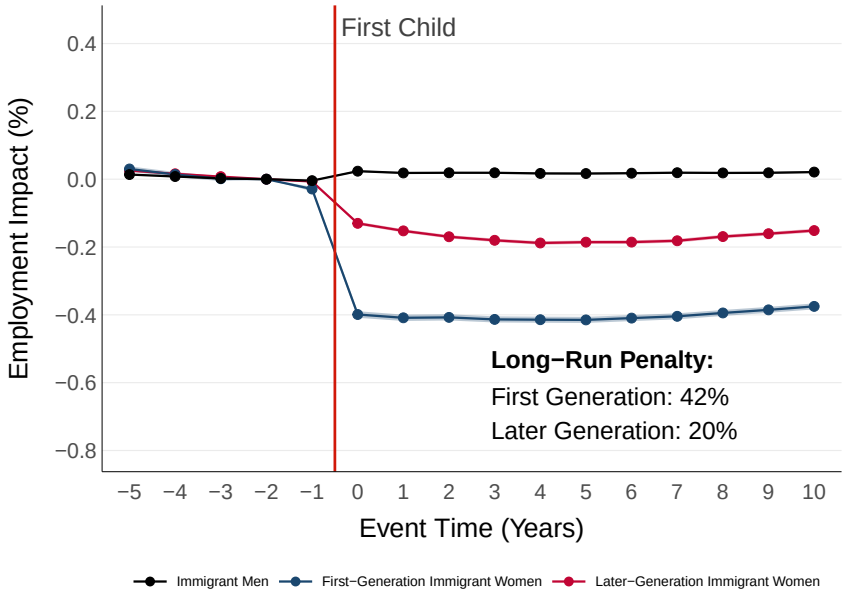
## FIGURE A.24: CULTURAL ASSIMILATION OF IMMIGRANTS

FIRST-GENERATION VS LATER-GENERATION CHILD PENALTIES BY ORIGIN-COUNTRY PENALTY

### A. Bottom Quartile of Child Penalty in Country of Origin




### B. Top Quartile of Child Penalty in Country of Origin




Notes: This figure presents event studies of first childbirth for first-generation and later-generation immigrants by quartile of the child penalty in country of origin. First-generation immigrants are defined as foreign-born US residents, while later-generation immigrants are defined as US-born residents who report foreign ancestry. The analysis is based on the 81 countries shown in Appendix Figure A.21, dividing countries into quartiles of the child penalty using the estimates in Kleven, Landais and Leite-Mariante (2025). The figure is constructed by running the event study specification (6) for first- and later-generation immigrant women separately (within the bottom and top quartiles of origin-country penalties, respectively). Each panel displays long-run child penalties (over event times 5-10) for first- and later-generation immigrants. The outcome is pooled employment (combining information on weekly and annual employment) and the sample is based on ACS data from 2000-2019 and CPS data from 1994-2020. The 95% confidence intervals are based on robust standard errors.