

Motherhood timing and the child penalty

Online appendix

Anikó Bíró*, Steven Dieterle[†], Andreas Steinhauer[‡]

*ELTE Centre for Economic and Regional Studies, Hungary. Email: biro.aniko@krtk.elte.hu. Anikó Bíró supported by the National Research, Development and Innovation Office of Hungary (grant number: OTKA K 146309). Anikó Bíró has no relevant or material financial interests related to the research described in this paper to disclose.

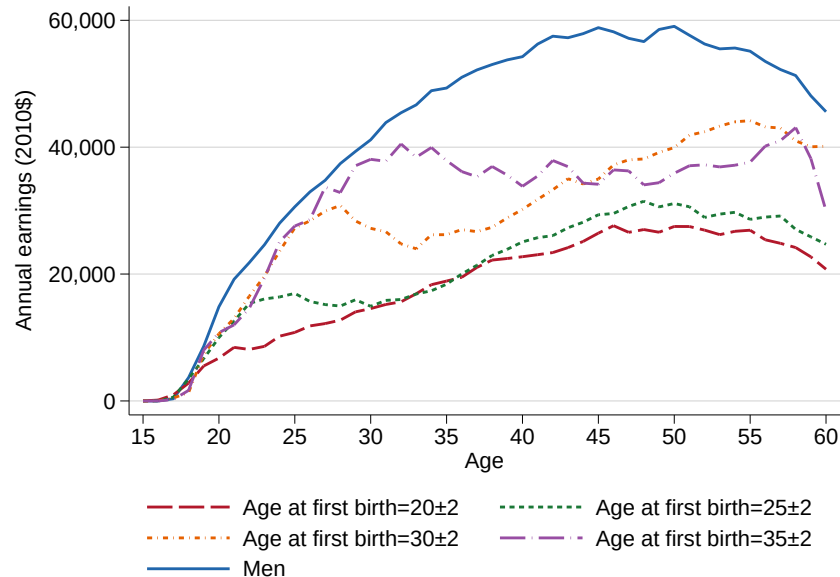
[†]The University of Edinburgh, UK. Email: steven.dieterle@ed.ac.uk. Steven Dieterle received no financial research funding support for the work and has no relevant or material financial interests related to the research described in this paper to disclose.

[‡]The University of Edinburgh, UK and CEPR. Email: andreas.steinhauer@ed.ac.uk. Andreas Steinhauer received no financial research funding support has no relevant or material financial interests related to the research described in this paper to disclose.

A US Results

In this Appendix, we provide baseline statistics and estimates from the US for the effect of the age at first birth on earnings and the child penalty. Here, we use data from the PSID (Panel Study of Income Dynamics).

Appendix Figure A1: Earnings and Age at First Birth in the US



Notes: Figure plots annual earnings constructed analogously to Panel (a) but based on the PSID. The sample is selected as in Kleven et al. (2019) and includes the years 1968-2015. Individuals who are not working are assigned zero earnings. Because of the switch to biennial sampling in 1997 we use moving 2-year average earnings after 1997. We convert annual earnings to real 2010\$ using the CPI. Due to sample size, we group women by ages at first birth 18-22, 23-27, 28-32, and 33-37.

Appendix Table A1: OLS Effect of Age at First Birth on Outcomes 5 Years After First Birth in the US

Dependent variable	Mean 5y After First Birth (1)	OLS Effect of One Year Delay (2)	Mechanical Effect of One Year Delay (3)	Child Penalty (4)	Effect on Child Penalty [95% CI] (5)
Annual earnings (\$)	18,561.9	1,437.5 (125.4)	919.7	-0.458	0.015 [0.008, 0.022]
<i>in %</i>		<i>0.077</i>			
Participation	0.594	0.010 (0.002)	0.009	-0.309	0.001 [-0.003, 0.005]
<i>in %</i>		<i>0.017</i>			

Notes: Annual earnings are total labour earnings per calendar year including zeroes. Participation is set to one if an individual ever worked in a calendar year, zero otherwise. We calculate means 5 calendar years after first birth and the OLS effect of age at first birth on the two outcomes 5 years after first birth, controlling only for year of birth of the mother (dummies). We divide the level effect by the means to show the effect in % in italics below robust standard errors in parentheses. The third column shows the mechanical effect of one year delay, which is calculated by estimating child penalties on the relevant sample, and then calculating what would happen if everyone delayed the average child penalty by one year according to equation (5). We show estimated child penalties 5 years after first birth in column (4) and the OLS effect from column (2) converted to an effect on the child penalty in column (5). For this we simply subtract the mechanical effect from the OLS effect and then divide by estimated counterfactual earnings. As a reading example for annual earnings: each year of delaying first birth in the U.S. is associated with a 1.5 percentage point smaller child penalty 5 years after first birth. The 95% CIs on those child penalty effects are shown in square brackets. See Section 4 for details on the estimation of the mechanical effect and the child penalty.

B Heterogeneous Treatment Effects

The arguments in Section 5 based on the OLS and IV estimates follows from the linear, homogeneous effects model presented in equation (1). However, in the presence of unmodeled heterogeneous effects both the OLS and IV estimates will produce different weighted averages of the underlying effects along the age at first birth distribution (Angrist and Pischke, 2009). Whether the interpretation holds up in the presence of heterogeneous effects depends on the true individual causal models relating age at first birth to the career outcomes we study, the relative weights placed on effects at different ages at first birth by the OLS and IV estimates, and the magnitude of the average biases for the two estimators. Our goal here will be to assess whether the alternative weighting implied by the OLS and IV estimators is likely to substantially change or threaten the comparisons we make based on the most likely sources of bias.

Obviously, we cannot fully characterize how a heterogeneous effects interpretation might affect the validity of our comparison of the estimators as we do not know the true individual causal models. However, it is possible to estimate the relative weights placed on effects at different ages at first birth by the OLS and IV estimators in order to assess how important the difference in weights might be for our interpretation. Start by considering a simplified version of our estimating equation excluding covariates to ease the exposition that follows:

$$y_i = \alpha + \beta A_i + \varepsilon_i \quad (\text{A1})$$

where y_i can represent any of our outcomes of interest and A_i is age at first birth. From here, we need to introduce notation for the underlying heterogeneous effects of delaying first birth for women. To do so, we follow Angrist and Pischke (2009) and define the individual causal model of potential outcomes for each woman indexed by i as

$$y_{Ai} = f_i(A)$$

with the derivative $f'_i(A)$ representing the individual causal effect of delaying age at first birth at different A . Next write the conditional expectation function (CEF) for y_i with respect to age at first birth as

$$h(a) \equiv E[y_i | A_i = a]$$

with derivative $h'(a)$. Note that $h'(a)$ is a descriptive object, simply how the expected value of the outcome changes with age at first birth — in order for this derivative to have a causal interpretation as the average of individual effects at a (that is for $h'(a) = E[f'_i(a)]$) would require A_i to be independent of y_{Ai} , an assumption that is at odds with our motivation for estimating bounds. However, the fact that this is not a causal object is helpful in the sense that we can flexibly estimate both $h(a)$ and $h'(a)$ nonparametrically without having to assume independence.

Angrist and Pischke (2009) show that the population regression of y_i on A_i will recover the following weighted average of the $h'(a)$ along the length of the age at first birth distribution:

$$\beta_{OLS} = \frac{\int h'(a) \mu_a^{OLS} dt}{\int \mu_a^{OLS} dt}$$

where

$$\mu_a^{OLS} = \{E[A_i | A_i \geq a] - E[A_i | A_i < a]\} \{P(A_i \geq a) [1 - P(A_i \geq a)]\} \quad (\text{A2})$$

are the OLS weights. The weights are determined by the distribution of A_i so we can easily

estimate them by replacing the expectations and probabilities in equation (A2) with their sample analogues. Therefore, we can separately estimate all the components that go into our OLS estimate of the returns to delay — both the weights and the derivatives of the CEF — without assuming independence.

Next we need to characterize the weighted average produced by our IV estimator — what Angrist and Pischke (2009) refer to as the Average Causal Response (ACR). For ease of exposition we assume that age at first birth takes on discrete, whole number values in the set $[\underline{A}, \underline{A} + 1, \dots, \bar{A} - 1, \bar{A}]$ where $\underline{A}$ and $\bar{A}$ are the minimum and maximum age at first birth in the population of interest. Further define the potential ages at first birth with the instrument, pregnancy loss, turned on or off as A_{1i} and A_{0i} , respectively. Under the typical assumptions needed for a Local Average Treatment Effect (LATE) interpretation of an IV estimand — namely that the instrument is independent of potential outcomes and treatments, the instrument can be excluded from the causal model, first stage relevance, and monotonicity (see Angrist and Pischke, 2009)— the IV estimand using our pregnancy loss instrument can be written as a weighted average of the average causal effect of a one unit increase in age at first birth at each a as follows:

$$\beta_{IV} = \sum_{a=\underline{A}}^{\bar{A}} \omega_a^{IV} E[Y_{ai} - Y_{a-1,i} | A_{1i} \geq a > A_{0i}]$$

where

$$\omega_a^{IV} = \frac{P[A_{1i} \geq a > A_{0i}]}{\sum_{j=\underline{A}}^{\bar{A}} P[A_{1i} \geq j > A_{0i}]}$$

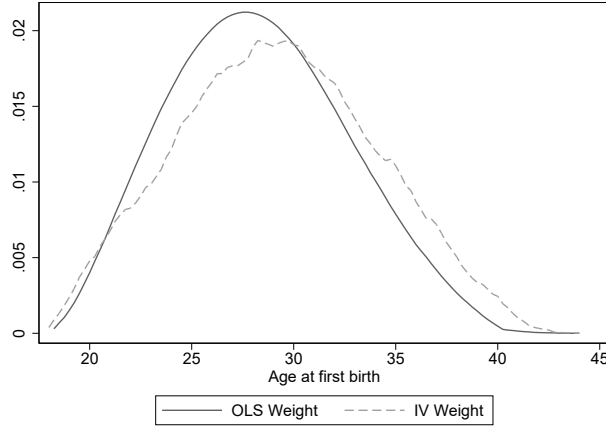
are the IV weights. Note, unlike the OLS estimand which was written as a weighted average of descriptive — not causal — parameters, the IV estimand is a weighted average of the unit causal effects that do have a causal interpretation. Therefore, without assuming instrument independence or exclusion we would not be able to directly estimate these parameters. The fact that we only have a binary instrument also limits the extent of heterogeneity we can account for. However, as noted by Angrist and Pischke (2009), we can estimate the IV weights under instrument independence as the difference in the Cumulative Distribution Function of age at first birth for those who suffer a loss and those who do not.

Figure A2 depicts the estimated weights for the OLS and IV estimators at different ages at first birth. As might be expected, the IV estimator places more weight on later ages at first birth due to the fact that pregnancy loss necessarily induces a delay.¹ Again, how

¹Note that we limit our sample to those with a first event — live birth or pregnancy loss — at age 40 or lower. Therefore, the OLS weights hit zero at ages above 40, but the IV weights can be non-zero above age 40 due to mothers who suffer a loss at first event and are delayed until after 40 for their first birth.

this shift in weight to later ages at first birth impacts our interpretation will depend on the shape of the underlying true average partial effects, $E[f'_i(a)]$, and the average biases of the two estimators. While we cannot directly assess this, we can perform the following thought exercise to explore how important the difference in weighting would be if the *only* difference between the OLS and the IV estimators were the different weights. That is, since we can estimate $h'(a)$ we can check how swapping ω_a^{IV} for μ_a^{OLS} in equation A2 changes the weighted average produced. If changing the weights considerably changes the average, then we need to be cautious in relying on arguments based on the typical or average factors that would bias the IV and OLS estimators of equation (A1). However, if the difference in weighting results in only minor changes to the average produced, then we have more confidence in our interpretation based on the most common potential sources of bias.

Appendix Figure A2: OLS and IV Weight Estimates

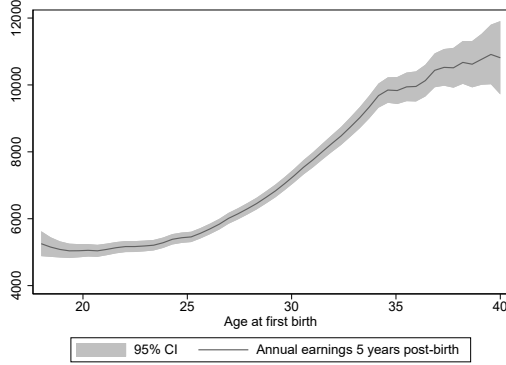


Notes: This figure plots the estimated weights placed on partial effects at different ages at first birth by the OLS and IV estimators with no additional control variables.

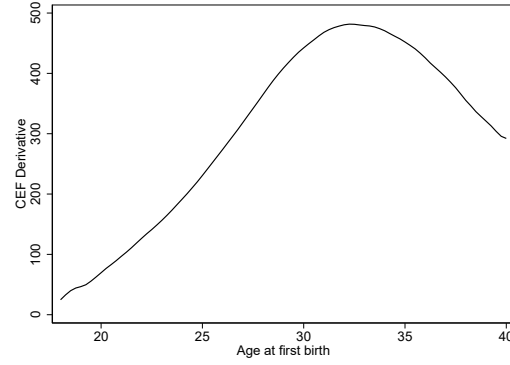
We, therefore, start by estimating $h(a)$ and $h'(a)$ nonparametrically for one of our key outcome measures: annual earnings five years after first birth. Figure A3 depicts the estimated CEF and derivatives of the CEF for this outcome. In Panel (A3a), we see the CEF for earnings five years post-birth is first increasing and convex, but has an inflection point at ages at first birth in the early thirties — resulting in an initially increasing and then decreasing derivative in Panel (A3b).

Appendix Figure A3: Nonparametric CEF and Derivative Estimates

(a) Annual Earnings Five Years Post-birth CEF



(b) Annual Earnings Five Years Post-birth CEF Derivative



Notes: This figure plots nonparametric regression estimates of the conditional expectation function and its derivatives with respect to age at first birth for annual earnings five years after first birth. Annual earnings are total social security earnings in the calendar year 5 years after first birth. Women who are not employed in that year are included with zero earnings.

With the estimated derivatives of the CEF — $\widehat{h'(a)}$ — and the estimated OLS and IV weights in hand, we can calculate the difference in the IV and OLS estimators due purely to the different weighting schemes. Table A2 reports the average of the CEF derivative estimates using the estimated IV and OLS weights — that is the mean of the derivatives depicted in Panel (A3b) of Figure (A3) using the weights depicted in Figure (A2). For reference, we also include the actual IV and OLS estimates of equation (A1). For earnings five years after first birth, our IV estimate with no covariates is 72% smaller than the OLS estimate. When weighted by the estimated OLS weights, the average $\widehat{h'(a)}$ is very close to the actual OLS estimate — an implied 330.43 euro increase in earnings from a single year of delay as opposed to a 334.82 change. This provides confidence that separately estimating the CEF derivatives and OLS weights matches the OLS estimator closely. When we use the IV weights the average derivative is slightly larger — 12.60 euros more, or only 3.8% larger — than the OLS weighted average derivative. The fact that these are so close is natural given the increasing then decreasing shape of the $\widehat{h'(a)}$ in Panel (A3b) of Figure (A3). This difference is also small relative to the raw difference in the OLS and IV estimates — suggesting that factors that bias the estimators away from uncovering weighted averages of the same true causal effects — such as the negative health effects of pregnancy loss or preferences for later motherhood being correlated with ability — are generally more important than the difference in weights.

Appendix Table A2: Weighted Average CEF Derivatives

Dependent variable	IV	OLS	IV Weighted CEF Derivatives	OLS Weighted CEF Derivatives
Annual Earnings 5 Years Post-birth	93.37	330.43	347.42	334.82

Notes: This table shows the estimated return to delay annual earnings 5 years after first birth estimated by OLS or IV with no controls as well as the mean CEF derivatives weighted by the estimated OLS and IV weights. Annual earnings are total social security earnings in the calendar year 5 years after first birth. Women who are not employed in that year are included with zero earnings. The CEF derivatives are estimated by nonparametric regression and the weights are estimated as outlined in Appendix B.

C Education Level at First Birth

The mother’s education information is recorded in the birth registry at the time of birth under the following categories: compulsory, some high school, high school, professional apprenticeship, teacher’s college, and college. We thus only have information about the highest education achieved for women who ever give birth (groups 1 and 2). Table A3 shows shares of education levels for women whose first event we observe is a live birth, and those whose first observed event is a pregnancy loss, followed eventually by a live birth.² We focus on the subsample of women with non-missing education information (about 75% of the main sample).

Importantly, since the education level is recorded at the time of the birth of a child, it reflects the investments in education made up to that point. This allows us to consider whether delaying motherhood may provide an opportunity to invest in human capital before having a child— at which point the opportunity costs of further investment may be higher. Given the categorical nature of the education variable, we model the probability of having each education level at the birth of the child by a multinomial logit (MNL):

$$P(E = j|A1B, x) = \frac{\exp(A1B_i\gamma_j + \mathbf{x}_i\beta_j)}{1 + \sum_{h=1}^J \exp(A1B_i\gamma_h + \mathbf{x}_i\beta_h)} \quad (\text{A3})$$

where $j = 1, \dots, J$ indicate the education level and $\mathbf{x}_i$ includes mother’s cohort indicators (rounded to 5 year bins) and the age at entry to the labour market. We set the compulsory education level as the base category with the other categories generally reflecting higher

²Unfortunately, data quality is an issue with many missing values: 23.8% among group 1 and 43.6% among group 2. The birth registry data ends in 2007, which explains the higher share of missings among women whose first birth is delayed by pregnancy loss.

educational investments and present estimated odds ratios for the age at first birth in Table A4. We also present the joint test of whether the coefficients on age at first birth are equal to zero for all education levels ($\gamma_j = 0 \forall j$)— or effectively whether the odds ratio relative to compulsory schooling is equal to one for each.

In column (1), we see that age at first birth is a significant predictor of education level— both for each category relative to compulsory school and jointly across category— with higher age at first birth associated with an increased probability of attaining higher educational credentials (odds ratio greater than one). While this relationship may be driven by a host of factors, it is consistent with human capital accumulation playing a role in the returns to delaying motherhood. In contrast, when we replace age at first birth with an indicator for suffering a pregnancy loss in column (2), the pregnancy loss indicator is not a statistically significant predictor of education level. This suggests that the sudden, surprise delay induced by the pregnancy loss is not likely to increase educational investments— indeed, the estimated odds ratios relative to compulsory schooling are all very close to or less than one. Therefore, using the pregnancy loss instrument for identification of delay effects may not capture an important mechanism for delay to improve labour market outcomes.

Appendix Table A4: Education Level at First Birth

Education Category	Odds Ratio	
	(1) Age at First Birth	(2) Pregnancy Loss
Professional Apprenticeship	1.0564 (0.0089)***	0.9064 (0.0610)
Some High School	1.0853 (0.0100)***	0.9853 (0.0735)
High School	1.1158 (0.0105)***	0.9173 (0.0716)
Teacher's College	1.105 (0.0141)***	0.9946 (0.1065)
College	1.1887 (0.0155***)	0.8902 (0.0984)
$H_0 : \gamma_j = 0 \forall j$	$\chi^2_5 = 240.52$ p-value=0.0000	$\chi^2_5 = 4.47$ p-value=0.4841

Notes: All columns include indicator variables for year of birth and age at labour market entry. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A3: Education by Group

	First event is live birth (G1)	First event is preg. loss, subsequent birth (G2)	Difference
Education: primary	0.1268 (0.3328)	0.1305 (0.3370)	-0.0037 (0.0067)
Education: apprenticeship	0.4400 (0.4964)	0.4117 (0.4922)	0.0284 (0.0101)***
Education: some HS	0.1755 (0.3804)	0.1884 (0.3911)	-0.0129 (0.0077)*
Education: HS	0.1529 (0.3599)	0.1532 (0.3603)	-0.0003 (0.0073)
Education: teachers college	0.0482 (0.2141)	0.0537 (0.2255)	-0.0056 (0.0044)
Education: college	0.0565 (0.2310)	0.0624 (0.2420)	-0.0059 (0.0047)
Observations	31,089	2,643	33,732

Notes: Women with missing education information (23.8% among group 1 and 43.6% among group 2) excluded. First two columns show means and standard deviations in parentheses. Last column shows difference in means between the groups and standard errors of the differences in parentheses. Differences tested with t-test allowing for unequal variance. *** p<0.01, ** p<0.05, * p<0.1.

D Inverse Probability Weighting Details

To start, denote the vector of data for an individual by $\mathbf{w}_i = (y_i, A_i, \mathbf{x}_i, z_i)$. Let our random draw from the population be given by $(\mathbf{w}_i, s_i)$ where s_i is a binary indicator of selection, where when $s_i = 0$ we only observe part of $\mathbf{w}_i$. In our case, $s_i = 0$ indicates women with a pregnancy loss and no subsequent observed birth—those for who we do not observe y_i or A_i —and $s_i = 1$ indicates all other women in our sample—both the no pregnancy loss women and those who suffer a pregnancy loss but are observed with a subsequent birth.

Our aim is to estimate the parameters of equation (1). Following Wooldridge (2010), denote $g(\mathbf{w}_i)$ as any scalar function of the data where the mean exists. Note that $E[g(\mathbf{w}_i)]$ could represent any of the moments used for identification for our estimators of interest. Instead of using data representative of the population $E[g(\mathbf{w}_i)]$, we only have the selected sample to identify $E[s_i g(\mathbf{w}_i)]$. The goal is to recover the population moment from the selected subsample.

The IPW solution requires us to have some variables that we observe for everyone ($s_i = 1$ and $s_i = 0$) that are good predictors of selection. Wooldridge (2010) sets out the following assumptions for IPW:

IPW1 $\mathbf{w}_i$ is observed when $s_i = 1$

IPW2 There is a vector $\mathbf{r}_i$ such that $P(s_i = 1 | \mathbf{w}_i, \mathbf{r}_i) = P(s_i = 1 | \mathbf{r}_i) \equiv p(\mathbf{r}_i)$

IPW3 For all $\mathbf{r} \in \mathcal{L} \subset \mathbb{R}^J$, $p(\mathbf{r}) > 0$

IPW4 $\mathbf{r}_i$ is observed when $s_i = 1$

IPW1 simply states that we observe all of $\mathbf{w}_i$ when women are observed to have a child. In practice, $\mathbf{r}_i$ will include pre-pregnancy health and labour market information so that IPW4 clearly holds in our setting as we observe this for everyone regardless of selection. IPW3 just states that there is no set of predictors for which the probability of selecting into the sample is zero—that is, no matter what someone’s pre-pregnancy loss earnings or health are they have some chance of having a subsequent child.

IPW2 is the key assumption—assuming that $\mathbf{r}_i$ is such a good predictor of selection that once we condition on it the selection probability is independent of the other elements of $\mathbf{w}_i$. IPW2 is an ignorability or unconfoundedness assumption for selection conditional on the variables included in $\mathbf{r}_i$.

In order to implement the IPW, we need to estimate the selection probabilities, $p(\mathbf{r}_i)$. To account for the fact that only women who suffer a pregnancy loss have the opportunity

to select out, we write $s_i = z_i + (1 - z_i)\tau_i$ where

$$\tau_i = \begin{cases} 0 & \text{if we do not observe a birth post-pregnancy loss} \\ 1 & \text{if we do observe a birth post-pregnancy loss.} \end{cases}$$

Since we can include z_i in $\mathbf{r}_i$, we have:

$$P(s_i = 1|\mathbf{r}_i) = E(s_i|\mathbf{r}_i) = z_i + (1 - z_i)E(\tau_i|\mathbf{r}_i)$$

Therefore, we need to estimate $E(\tau_i|\mathbf{r}_i) = P(\tau_i = 1|\mathbf{r}_i) = p_\tau(\mathbf{r}_i)$, and then plug it into our expression for s_i . In practice we estimate the probabilities by probit using a rich set of pre-event labour market — effectively pre-pregnancy lags of our key outcome variables — and health variables. By including pre-pregnancy labour market variables, we can control for many of the unobservable factors that may affect post-pregnancy labour market outcomes when accounting for the post-loss selection. This is conceptually similar to the arguments in favor of using pre-treatment dependent variables when creating synthetic control groups (Abadie et al., 2010).

Given the probit fitted values $\hat{p}_\tau(\mathbf{r}_i)$ we can write the estimated selection probabilities as:

$$\hat{p}(\mathbf{r}_i) = z_i + (1 - z_i)\hat{p}_\tau(\mathbf{r}_i) \tag{A4}$$

and use $1/\hat{p}(\mathbf{r}_i)$ as the weights in our regressions.

We model the probability of being observed with a birth after a pregnancy loss, $p_\tau(\mathbf{r}_i)$, as a probit where $\mathbf{r}_i$ includes:

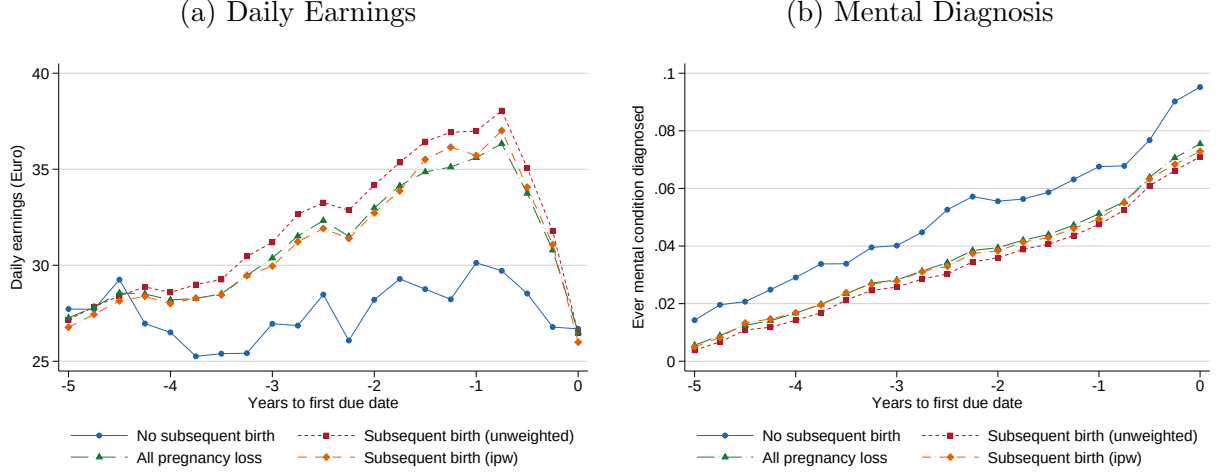
1. Indicators for the mother's birth cohort and year of the pregnancy loss
2. Pre-pregnancy annual histories: employment, earnings, health care utilization, depression diagnosis and medication use, psychological diagnosis and medication use

We then use the estimated probabilities, $\hat{p}_\tau(\mathbf{r}_i)$ to generate our IPWs as $1/\hat{p}(\mathbf{r}_i)$ where $\hat{p}(\mathbf{r}_i) = z_i + (1 - z_i)\hat{p}_\tau(\mathbf{r}_i)$.

Figure A4 provides evidence on how reweighting by the IPW may help account for the positive selection into the pregnancy loss sample used for estimation. Panel (a) presents four separate series of daily earnings in the five years before the first pregnancy for women who suffer a pregnancy loss. Again we see the positive selection into the estimation sample with the unweighted earnings being higher for the subsequent birth sample than for the entire pregnancy loss group. Importantly, when we reweight the subsequent birth sample using

the IPW, the series overlaps the pre-pregnancy earnings for the entire pregnancy loss group. Panel (b) shows a similar pattern for the mental health diagnoses before pregnancy. This suggests that weighting the subsequent birth group better reflects the experience of those that suffer a pregnancy loss— including those that will fall out of the natural experiment— thereby better matching the thought experiment behind the pregnancy loss IV approach.

Appendix Figure A4: Earnings and mental health diagnoses with and without IPW



Finally note that our IPW approach aims to recover the effect of delay for the population of women who would be observed having a pregnancy in our data — those with an observed live birth or pregnancy loss. This encompasses all women who we observe with a child in the no loss group — a mixture of the two latent types indexed by τ . Importantly, this includes those who have a child because they do not suffer a loss ($z_i = 1$), but would not be observed with a child had they suffered a loss ($\tau_i = 0$). Put differently, we wish to estimate the treatment effect across all types of women we see with a child in the no-loss group. An alternative approach to the selection problem would be to focus on the population of women with $\tau_i = 1$ only, i.e., those who would “select” into the sample if they suffer a loss — more akin to a treatment on the treated effect for those with $\tau_i = 1$. Note that this treatment on the treated type effect would not be recovered by the unweighted estimates, rather, it would require using probability — rather than inverse probability — weights based on an out-of-sample prediction for the no loss group using the estimated coefficients of $p_\tau(\mathbf{r}_i)$ from the loss group. While this may better reflect the effect for those who are more able to or have a stronger desire to have children, it would overlook the effect for those with $\tau_i = 1$, but are observed with a child since they do not suffer a loss ($z_i = 1$). We prefer our approach that better captures the effects for all types of women we see with children. Table A9 compares the results using our IPW approach, unweighted estimates, and using a modified

weighting approach using probability weights $p_\tau(\mathbf{r}_i)$. Overall both weighting approaches yield quantitatively similar results.

E Censored Child Penalty and Cumulative Earnings

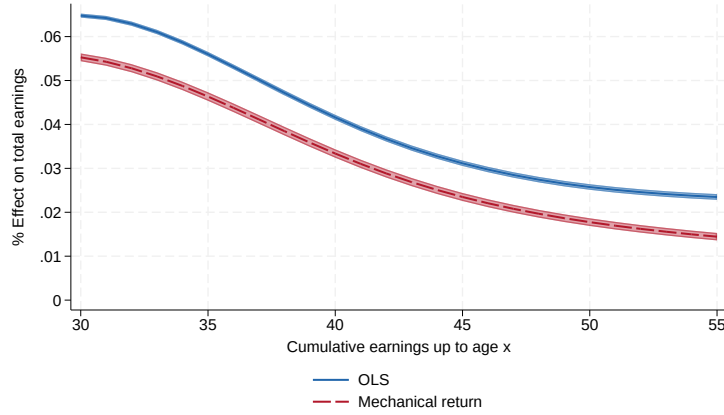
In this appendix we consider the importance of censoring when estimating the effect of delay on cumulative outcomes. We do this to demonstrate why our estimates based on the cohort of women whom we observe from ages 15-55 are much smaller than previous literature³ and why we cannot use the censored health data sample to estimate career returns to delay. However, we can still compare the estimated returns to simulated returns holding the child penalty constant. In line with the main text results, we find that the mechanical effect of delaying a fixed penalty comprises the majority of the effect of delay in the censored case.

Censoring arises because we observe the women in our health data sample up to 2017, when the youngest are only age 29 (the oldest are 58). Intuitively, if a large part of the return to delay comes from higher total pre-birth earnings (more years on the childless profile) instead of altering post-birth earnings, then the gains from delaying first birth are a bigger and bigger part of total earnings as the age window gets smaller.

The simplest way to show the effect of censoring is to take the 1957-1962 cohort for which we observe the entire labour market career and progressively lower the age up to which cumulative earnings are calculated (age 54, then 53, and so on). At each step we estimate the OLS return to delay regressing total earnings up to e.g., age 50 on age at first birth and year of birth and labour market entry dummies (according to equation (1)). Next, we simulate the effect of delaying a fixed relative child penalty and calculating total earnings with and without delaying first birth also up to age 50, which gives us the “mechanical effect” of delaying first birth by a year. The procedure is exactly the same as in Table 5, but done separately in yearly steps from age 30 to age 55. The resulting estimates are shown in Figure A5. As we lower the censoring age for earnings, both the OLS and mechanical returns estimates tend to increase considerably — consistent with the concerns outlined above — with mechanical and OLS returns above 5% and 6%, respectively, when we measure earnings up until women are in their early-thirties. However, note that at each step the mechanical return comprises the majority of the overall return to delay as estimated by OLS. The censoring problem does not change the result that the majority of the effect of delaying first birth comes from delaying a fixed penalty.

³e.g., Miller (2011) finds returns of 9.63% and 6-8% in the U.S. using OLS and IV, respectively, by limiting the age range to 21-33.

Appendix Figure A5: Censoring and the Career Returns of Delay

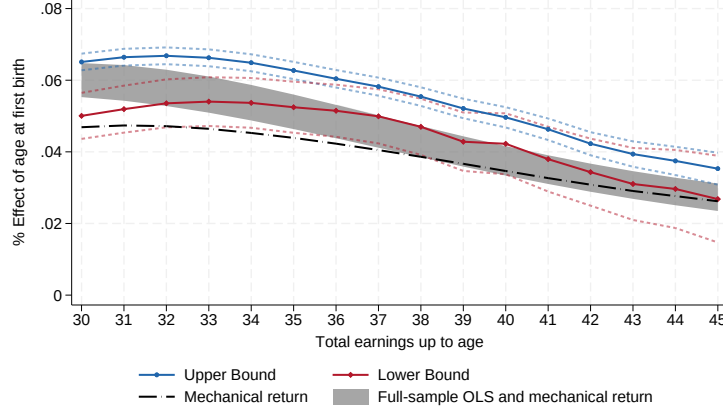


Notes: Figure plots the OLS and mechanical returns to delay on cumulative earnings estimated as in Table 5 on the cohort of Austrian women born 1957-1962, whose labour market careers we observe almost entirely (age 15-55, after which retirement starts, cf Figure 1). We calculate their total earnings up to age 30, 31, ..., 55, and estimate the OLS return to delaying first birth for total earnings up to this age. We then use the long-run child penalties to simulate their earnings had they suffered the fixed average child penalty that is independent of age at first birth. We then regress those simulated earnings on age at first birth to get the mechanical return to delay. Throughout we include the same controls as in our main specifications (cohort and age at labour market entry dummies) but we include women from all of Austria, not just those we can link to health insurance data. Estimated long-run penalties go up to 20 years after first birth and we keep them fixed at the 20-year value for 21+ years after first birth. Long-run penalties have been declining in Austria (see Kleven et al., 2024) so we only use women born 1957-1967 to estimate the long-run penalties used in this figure.

We next consider how the results for the 1957-1962 cohort for all of Austria compare to the sample from Upper Austria that can be linked to the pregnancy loss information. Figure A6 plots our estimated IV and NR estimates using the IPW and converted into percentages by dividing by averages along with the mechanical return for total cumulative earnings for the group with health data available. We start by measuring earnings up to age 45, which can be done for women born between 1959 and 1972. We then proceed to re-estimate the returns after lowering the censoring age one year at a time (measure earnings up to age 44, then 43, etc.) and adding in successive cohorts (1973, then 1974, etc.) until we use a measure of earnings up to age 29 for all women in our main sample.⁴ Therefore, as we move along the curves in the figure we are changing both the censoring age and the cohort composition for the sample. The shaded region in the figure depicts the corresponding OLS and mechanical returns for the sample from all of Austria — again changing both the censoring age and the cohort composition at each point.

⁴As we look at earlier birth cohorts this introduces some degree of selection since we only consider later ages at first birth that fall within our health data window after 1999.

Appendix Figure A6: Estimates on Causal Effect of Delay on Total Earnings



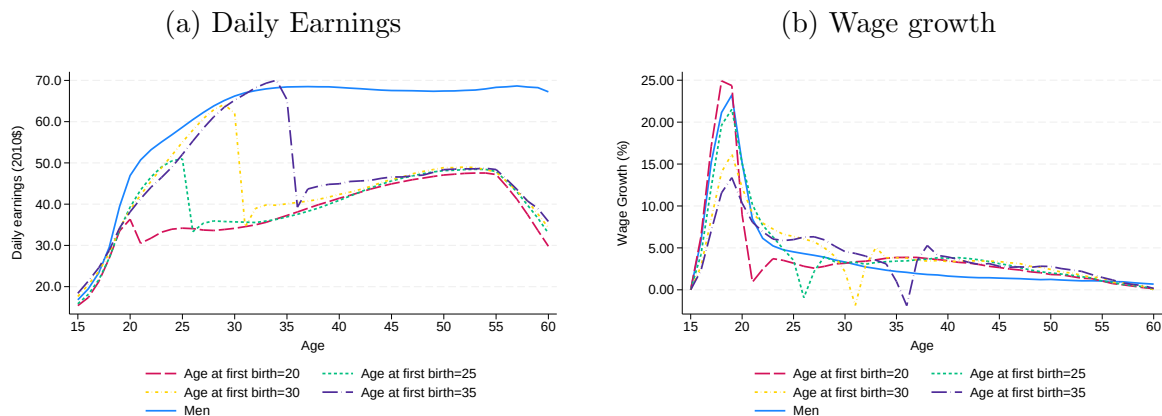
Notes: This figure plots the estimated IV and NR estimates for the return to delay on total earnings from age 15 to x . These estimates are calculated using our IPW weights. We also add the mechanical return to delay by predicting total earnings based on the child penalty regression (2) estimated for the mothers in our sample and then regressing predicted total earnings from age 15 to x on age at first birth. For $x = 30$ we include cohorts born 1959-1987. For each additional year of age we exclude one cohort of women who are not observed at that age in 2017 (the last year in our data). e.g., for total earnings over ages 15 to 45 we include women born 1959-1972. The shaded area shows the OLS and mechanical return from Figure A5 for comparison.

Censoring earnings at age 45, we estimate the return to a one year delay to lie between 2.68% and 3.53%. As with the 1957-1962 cohort, lowering the censoring age tends to increase the estimated returns. Interestingly, while the mechanical return estimates are slightly smaller than the IV estimates, they do track the IV estimates closely as we change the censoring age — coinciding almost exactly when we measure earnings up to age 45, with an estimated mechanical return of 2.62%. Importantly, the OLS/mechanical returns for the 1957-1962 cohort also follow the estimates based on our main sample closely. This suggests that the mechanical return estimates reflect a reasonable lower bound to pair with the OLS upper bound yielding bounds for the career earnings return to a one year delay of motherhood for the 1957-1962 cohort of between 1.4% to 2.3% — considerably smaller than most other estimates in the literature once we take into account both the censoring and post-loss selection issues.⁵

⁵This also suggests that the returns to delay for the sample from Upper Austria are fairly representative of the returns for women across Austria.

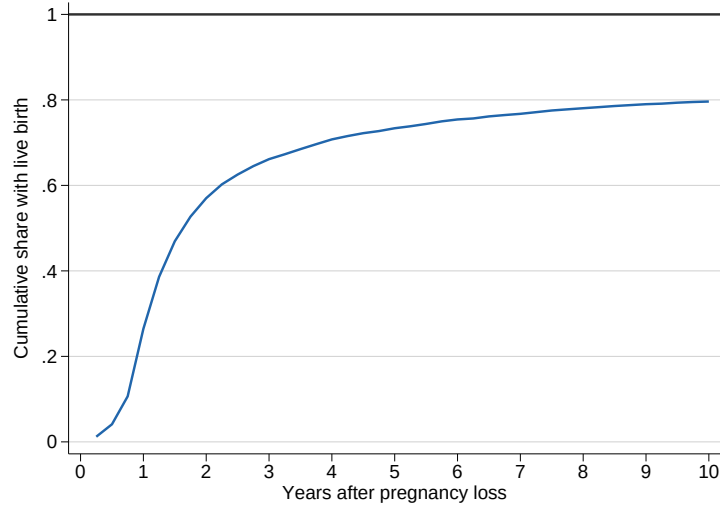
F Additional Figures and Tables

Appendix Figure A7: Daily Earnings, Wage Growth and Age at First Birth



Notes: Panel (a) plots daily social security earnings by age converted to 2010 euro (CPI) and then to USD using the ECB reference exchange rate. Panel (b) shows wage growth in percentage terms. The sample for this figure comprises every Austrian observed in the social security data (all of Austria) between 1972-2017. Years without earnings are included as zeroes. Years after the person has died are coded as missing. All Austrian men are included. We assign age at first birth based on year of the first observed birth spell minus year of birth and separately plot annual earnings by age for ages at first birth 20, 25, 30, and 35. We correct for topcoding using a simple Pareto distribution estimated by age and gender on income tax data (1994-2012).

Appendix Figure A8: Cumulative Share of First Births After Pregnancy Loss



Notes: The sample in this figure is all women who we observe with a pregnancy loss in 1999-2007 (see Section 3). We plot the share of this group who is ever observed with a live birth by year after the pregnancy loss event up to 10 years after loss. This ensures a balanced sample since we observe births up to and including 2017.

Appendix Table A5: Education and Distribution of Age at First Birth

Age at First Birth:	First Quartile	Median	Third Quartile
Only mandatory schooling	21.50	25.25	30.25
Professional apprenticeship	22.25	25.50	29.00
Some high school	24.25	27.25	30.75
High school	25.50	28.00	31.00
Teacher's college	26.25	28.50	31.25
College	29.00	31.50	34.00
All	23.5	27.00	30.50

Notes: Table shows the first quartile, median, and third quartile of age at first birth for each education group. Education is measured at the time of first birth, as is recorded in the birth register (missing for around third of the sample).

Appendix Table A6: Descriptive Statistics

	Means and std devs			Differences and std errs		
	First event is live birth (G1)	First event is pregnancy loss		G1-G2	G1-G3	G2-G3
		Subsequent birth (G2)	No subsequent birth (G3)			
Age at first event	27.332 (4.746)	27.317 (5.060)	31.739 (5.738)	0.015 (0.074)	-4.407 (0.147)***	-4.422 (0.176)***
Age at first birth	27.332 (4.746)	29.561 (5.222)		-2.229 (0.074)***		
Total number of births	1.915 (0.802)	1.846 (0.781)		0.069 (0.012)***		
Average daily earnings (Euro)	41.769 (15.665)	43.408 (15.274)	50.923 (18.755)	-1.639 (0.243)***	-9.155 (0.487)***	-7.515 (0.543)***
Experience in years	17.031 (6.688)	17.459 (6.735)	21.435 (9.318)	-0.428 (0.104)***	-4.403 (0.209)***	-3.976 (0.248)***
Employed before event	0.870 (0.336)	0.889 (0.314)	0.833 (0.373)	-0.019 (0.005)***	0.037 (0.010)***	0.056 (0.011)***
Daily earnings before event (Euro)	45.903 (27.708)	47.474 (27.208)	43.990 (30.225)	-1.571 (0.429)***	1.913 (0.858)**	3.484 (0.944)***
Age at labour market entry	18.110 (3.642)	18.168 (3.809)	20.130 (5.868)	-0.058 (0.057)	-2.020 (0.115)***	-1.962 (0.145)***
Log total healthcare utilization	8.128 (1.264)	8.511 (0.704)	8.177 (1.201)	-0.382 (0.019)***	-0.048 (0.039)	0.334 (0.028)***
Log total healthcare utilization before event	5.131 (2.720)	5.693 (2.429)	5.432 (2.837)	-0.562 (0.042)***	-0.302 (0.084)***	0.260 (0.085)***
Observations	46,437	4,561	1,071	50,998	47,508	5,632

Notes: Group 1—first event is live birth. Group 2—first event is pregnancy loss, subsequent live birth is observed. Group 3—first event is pregnancy loss, subsequent birth is not observed. All earnings and utilization values in real (CPI base 2000) euro.

First three columns show means and standard deviations in parentheses. Last three columns show differences in means between the groups and standard errors of the differences in parentheses. Significance tests against zero with t-test allowing for unequal variance. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A7: Effects on Child Penalties in Annual Earnings by Years After First Birth

Years from first birth	Level Effects			Mechanical Effect of One Year Delay	Child Penalty	Effects on Child Penalty (Relative to Mechanical Return)		
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{NR}^{IPW}$	$\hat{\beta}_{IV}^{IPW}$			$\hat{\beta}_{OLS}$	$\hat{\beta}_{NR}^{IPW}$	$\hat{\beta}_{IV}^{IPW}$
$t = 0$	366.3 (12.2)	234.1 (17.3)	-146.9 (50.2)	192.4	-0.645	0.010 (0.001)	0.002 (0.001)	-0.019 (0.003)
$t = 1$	70.1 (7.7)	40.9 (10.7)	-49.7 (30.5)	30.9	-0.932	0.002 (0.000)	0.001 (0.001)	-0.004 (0.002)
$t = 2$	35.3 (12.2)	-6.7 (16.4)	-150.6 (45.7)	69.1	-0.824	-0.002 (0.001)	-0.004 (0.001)	-0.012 (0.002)
$t = 3$	203.4 (14.3)	127.3 (19.4)	-126.6 (56.9)	94.6	-0.733	0.006 (0.001)	0.002 (0.001)	-0.011 (0.003)
$t = 4$	256.7 (15.3)	187.8 (21.3)	-48.8 (64.0)	98.9	-0.703	0.008 (0.001)	0.005 (0.001)	-0.007 (0.003)
$t = 5$	307.6 (16.1)	237.3 (22.6)	-5.2 (69.6)	106.9	-0.669	0.010 (0.001)	0.007 (0.001)	-0.006 (0.003)
$t = 6$	356.7 (16.6)	285.0 (23.3)	42.5 (73.6)	117.8	-0.632	0.012 (0.001)	0.008 (0.001)	-0.004 (0.004)
$t = 7$	391.6 (17.1)	327.6 (23.9)	107.4 (78.6)	127.0	-0.605	0.013 (0.001)	0.010 (0.001)	-0.001 (0.004)
$t = 8$	413.3 (17.7)	345.1 (25.1)	105.9 (86.3)	136.9	-0.581	0.013 (0.001)	0.010 (0.001)	-0.001 (0.004)
$t = 9$	451.4 (18.5)	369.2 (26.8)	78.8 (96.3)	144.4	-0.560	0.014 (0.001)	0.010 (0.001)	-0.003 (0.004)
$t = 10$	488.2 (19.4)	383.2 (36.7)	-19.3 (140.5)	150.2	-0.541	0.015 (0.001)	0.011 (0.002)	-0.008 (0.006)

Notes: See notes to Tables 2 and 3 for explanations of variables and methods. This table shows the OLS, Nevo-Rosen (NR) and IV estimates. The Nevo-Rosen and IV estimates are from the specifications using IPW. Last three columns show the coefficients minus the mechanical effect divided by counterfactual earnings. This translates the level effects into effects on the child penalty over and above what we would expect to see from mechanically being one year older when delaying first birth. Robust standard errors in parentheses.

Appendix Table A8: Outcomes 5 Years After First Birth (Robustness)

Dependent variable	Unweighted		IPW		
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{IV}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{NR}$	$\hat{\beta}_{IV}$
Panel A. Excluding Age at Labour Market Entry Dummies					
Annual earnings (€)	381.6 (16.0)	57.3 (69.7)	371.2 (16.5)	291.4 (22.4)	-19.1 (69.5)
Days employed	5.156 (0.298)	2.873 (1.356)	5.078 (0.304)	4.396 (0.415)	1.741 (1.328)
Daily earnings (€)	0.767 (0.154)	0.335 (0.680)	0.723 (0.151)	0.582 (0.156)	0.060 (0.573)
Employed	0.009 (0.001)	0.005 (0.004)	0.009 (0.001)	0.008 (0.001)	0.002 (0.004)
White-collar job	0.021 (0.001)	-0.000 (0.005)	0.019 (0.001)	0.014 (0.001)	-0.007 (0.005)
Back to pre-birth firm	0.019 (0.001)	0.015 (0.004)	0.019 (0.001)	0.017 (0.001)	0.009 (0.004)
Panel B. Including Probit Controls					
Annual earnings (€)	159.6 (46.0)	-31.4 (63.9)	172.5 (47.8)	86.1 (50.9)	-42.1 (66.4)
Days employed	2.219 (0.826)	1.990 (1.268)	2.539 (0.841)	2.158 (0.945)	1.593 (1.314)
Daily earnings (€)	0.279 (0.265)	0.135 (0.661)	0.234 (0.257)	0.132 (0.344)	-0.022 (0.602)
Employed	0.003 (0.002)	0.002 (0.004)	0.003 (0.002)	0.002 (0.003)	0.001 (0.004)
White-collar job	0.006 (0.003)	-0.005 (0.004)	0.005 (0.003)	0.001 (0.003)	-0.006 (0.004)
Back to pre-birth firm	0.010 (0.003)	0.008 (0.004)	0.011 (0.003)	0.010 (0.003)	0.010 (0.004)

Notes: See notes to Table 2 for explanation of table entries. The only change in this table is that we exclude age at labour market entry dummies in the regression (Panel A) or additionally include the Probit controls used to estimate the IPW (Panel B).

Appendix Table A9: Effect of Age at First Birth on Outcomes 5 Years After First Birth
(Different Weighting Method)

Dependent variable	IPW			PW		
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{NR}$	$\hat{\beta}_{IV}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{NR}$	$\hat{\beta}_{IV}$
Annual earnings (€)	300.5 (16.6)	237.3 (22.6)	-5.2 (69.6)	312.5 (16.1)	250.1 (21.2)	6.0 (65.9)
Days employed	5.055 (0.312)	4.399 (0.425)	1.884 (1.339)	5.179 (0.308)	4.607 (0.405)	2.370 (1.293)
Daily earnings (€)	0.460 (0.161)	0.398 (0.165)	0.169 (0.575)	0.499 (0.170)	0.447 (0.178)	0.258 (0.653)
Employed	0.009 (0.001)	0.007 (0.001)	0.003 (0.004)	0.009 (0.001)	0.008 (0.001)	0.004 (0.004)
White-collar job	0.020 (0.001)	0.015 (0.001)	-0.003 (0.005)	0.021 (0.001)	0.017 (0.001)	-0.002 (0.004)
Back to pre-birth firm	0.017 (0.001)	0.016 (0.001)	0.010 (0.004)	0.019 (0.001)	0.018 (0.001)	0.011 (0.004)

Notes: See notes to Table 2 for explanation of table entries. First three columns (IPW) replicate what is shown in Table 2. Last three columns (PW) show results of an alternative weighting method (probability weights – PW) discussed in Appendix D.

References

- Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, 2010.
- Joshua Angrist and Jörn–Steffen Pischke. *Mostly Harmless Econometrics: An Empiricists’ Companion*. Princeton Univerisity Press, Princeton, NJ, 2009.
- Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. Child penalties across countries: Evidence and explanations. *American Economic Association Papers and Proceedings*, 109:122–126, 2019.
- Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation. *American Economic Journal: Economic Policy*, 16(2):110–149, May 2024. ISSN 1945-7731. doi: 10.1257/pol.20210346. URL <https://www.aeaweb.org/articles?id=10.1257/pol.20210346>.
- Amalia R Miller. The effects of motherhood timing on career path. *Journal of Population Economics*, 24(3):1071–1100, 2011.
- Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.